

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

CALCUL DE LA VRAISEMBLANCE D'UNE COPULE BASÉE
SUR LES RANGS

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN MATHÉMATIQUES
CONCENTRATION STATISTIQUES

PAR
ABDELILAH MAHFOUD

AOÛT 2017

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je voudrais tout d'abord adresser toute ma gratitude à mon directeur de recherche, Simon Guillotte, pour sa patience, sa disponibilité, son aide et surtout, ses judicieux conseils, qui ont contribué à alimenter ma réflexion.

Je remercie de tout mon cœur mes chers parents et ma famille qui ont toujours été là pour moi.

Je désire aussi remercier mes autres professeurs de maîtrise, Sorana, Juli et Karim, qui m'ont fourni les outils nécessaires à la réussite de mes études universitaires.

Enfin, je voudrais exprimer ma reconnaissance envers mes amis, Mahmoud, Raja, Mahfoud et Salim. Je vous remercie énormément pour vos encouragements, votre aide et votre grand soutien.

TABLE DES MATIÈRES

LISTE DES TABLEAUX	vii
LISTE DES FIGURES	ix
RÉSUMÉ	xi
INTRODUCTION	1
CHAPITRE I	
LES COPULES	3
1.1 Introduction	3
1.2 Fonction de répartition	3
1.3 Copules	4
1.4 Théorème de Sklar	5
1.5 Densité d'une copule	6
1.6 Exemples de copules	7
1.6.1 Copule d'indépendance (copule produit)	7
1.6.2 Copule gaussienne	7
1.6.3 Copule archimédienne	8
1.6.4 Copules polynomiales	9
1.7 Simulation	10
1.7.1 Génération d'observations (u, v) du vecteur aléatoire (U, V) d'une copule C	10
1.7.2 Génération d'observations (u, v) du vecteur aléatoire (U, V) dont la copule est de Clayton	11
1.7.3 Génération d'observations (x_1, x_2) du vecteur aléatoire (X_1, X_2) dont la fonction de répartition F a les marges continues	12
CHAPITRE II	
LA VRAISEMBLANCE BASÉE SUR LES RANGS	15

2.1	Introduction	15
2.2	Permutation	15
2.3	Statistiques du rang	17
2.4	Énoncé du problème	18
2.5	La vraisemblance des rangs	18
2.5.1	Exprimer la vraisemblance basée sur les rangs par une espérance	18
2.5.2	Exprimer la vraisemblance par une intégrale sur un simplexe	20
CHAPITRE III		
MÉTHODES UTILISÉES POUR APPROXIMER LA VRAISEMBLANCE		
	BASÉE SUR LES RANG	23
3.1	Méthode de Grundmann-Möller	23
3.1.1	L'intégrale d'un monôme sur un simplexe de dimension n	23
3.1.2	La formule d'intégration pour un simplexe de dimension n	26
3.2	Méthode de Monte-Carlo	29
3.2.1	Introduction	29
3.2.2	Principe de la méthode de Monte-Carlo	30
3.2.3	Application de la méthode de Monte-Carlo au calcul d'une intégrale sur $[0, 1]^n$	31
3.2.4	Simulation des variables aléatoires	34
3.2.5	Inversion de la fonction de répartition	35
3.3	Méthode de quasi-Monte-Carlo	36
3.3.1	Introduction	36
3.3.2	La discrépance	39
3.3.3	Calcul de la discrépance d'une séquence de points	40
3.3.4	Exemples des suites de points à discrépance faible	42

3.3.5	Théoreme de Koksma et Hlawka	44
CHAPITRE IV		
	COMPARAISON ENTRE LES TROIS MÉTHODES DANS LE CALCUL DE LA VRAISEMBLANCE BASÉE SUR LES RANGS	47
4.1	Introduction	47
4.2	Calcul de l'approximation de la vraisemblance basée sur les rangs avec les trois méthodes	47
4.2.1	Méthode de Monte-Carlo	47
4.2.2	Méthode de quasi-Monte-Carlo	49
4.2.3	Méthode de Grundmann-Möller	51
4.3	Comparaison entre les trois méthodes	53
4.3.1	La valeur exacte de la vraisemblance via la méthode de Grundmann- Möller	53
4.3.2	L'approximation par la méthode de Monte-Carlo	57
4.3.3	Comparaison entre les deux méthodes Monte-Carlo et quasi- Monte-Carlo selon la taille de l'échantillon	61
4.3.4	Comparaison entre les trois méthodes selon la taille de la per- mutation	65
	CONCLUSION	71
APPENDICE A		
	PREUVE DE LA PROPOSITION 2.5.1	73
APPENDICE B		
	PREUVE DE LA PROPOSITION (2.5.2)	75

LISTE DES TABLEAUX

Tableau	Page
4.1 Calcul de la vraisemblance $n!P(S(X, Y) = s)$ de la copule polynomiale de degré deux : $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ par la méthode de Grundmann-Möller pour $n = 4$ et 6 , et $\theta = 0, 8$	54
4.2 Calcul de la vraisemblance $n!P(S(X, Y) = s)$ de la copule $C_{a,b}(u, v) = uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b]$ avec la méthode de Grundmann-Möller, avec $a = -1.2$ et $b = 0.6$	56
4.3 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ tel que $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ par la méthode de Monte-Carlo, selon la taille de l'échantillon N . Pour un $\theta = 0.8$ et $n = 4$	58
4.4 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ selon la taille de l'échantillon N , par la méthode de Monte-Carlo, tel que $C_{a,b}(u, v) = uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b]$, où $a = -1.2$, $b = 0.6$ et $n = 4$	59
4.5 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ selon N la taille des échantillons utilisés dans les deux méthodes Monte-Carlo et quasi-Monte-Carlo. On a $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$, $\theta = 0.8$ et $n = 4$	62
4.6 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ de la copule $C_{a,b}(u, v) = uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b]$ avec les trois méthodes. Pour $a = -1.2$ et $b = 0.6$	63
4.7 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ selon la taille de la permutation n , par les trois méthodes, où $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$, $\theta = 0.8$, et la taille de l'échantillon $N_{MC} = N_{qMC} = 3000$	66
4.8 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ selon n la taille de la permutation s , par les trois méthodes. La copule $C_{a,b}(u, v) = uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b]$, $a = -1.2$, $b = 0.6$ et $N_{MC} = N_{qMC} = 3000$	67

LISTE DES FIGURES

Figure		Page
1.1	Génération d'observations (u, v) de la copule Clayton, pour $\theta = 7$.	12
3.1	À gauche : 100 points aléatoires sur $[0, 1]^2$. À droite : une grille de 10 par 10	37
3.2	Une séquence de $N = 20$ points et un intervalle (rectangle) $P \in \mathcal{J}_2^*$ dans le carré unité $[0, 1]^2$	39
3.3	Séquence de $N = 10$, points uniformément distribués sur $[0, 1]$. . .	41
3.4	Les premiers 256 points de la séquence de Halton en base 2 et 3. Les 10 premiers points de la séquence sont en rouge de 1 à 10, suivis par les points bleus de 11 à 100 ainsi que les points verts de 101 à 256. Figure tirée à partir de https://en.wikipedia.org/wiki/Halton_sequence 44	
4.1	La valeur de la vraisemblance $n!P(S(X, Y) = s)$ de la copule polynomiale de degré deux : $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ selon le degré d	54
4.2	L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ selon d avec $C_{a,b}(u, v) = uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b]$ et $a = -1.2$ et $b = 0.6$	57
4.3	La variation de l'approximation $n!P(S(X, Y) = s)$ selon la taille de l'échantillon N , par rapport à la valeur exacte <i>vr.GM</i> . Tel que $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$, $\theta = 0.8$ et $n = 4$	60
4.4	La variation de l'approximation $n!P(S(X, Y) = s)$ selon la taille de l'échantillon N , par rapport à la valeur exacte <i>vr.GM</i> $uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b]$, $n = 4$, $a = -1.2$ et $b = 0.6$	60
4.5	La variation de l'approximation de la vraisemblance $n!P(S(X, Y) = s)$ par les deux méthodes Monte-Carlo et quasi-Monte-Carlo selon la taille de l'échantillon N . Avec $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ avec $\theta = 0.8$ et $n = 4$	63

- 4.6 La variation de l'approximation de la vraisemblance $n! P(S(X, Y) = s)$ de la copule $C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$ selon la taille de l'échantillon N par les deux méthodes. Avec $a = -1.2$, $b = 0.6$ et $n = 4$ 64
- 4.7 Pourcentage d'erreur de l'approximation de la vraisemblance $n! P(S(X, Y) = s)$ selon n , pour les deux méthodes, où $C_\theta(u, v) = uv(1+\theta(1-u)(1-v))$, $\theta = 0, 8$ et $N_{MC} = N_{qMC} = 3000$ 68
- 4.8 Pourcentage d'erreur de l'approximation $n! P(S(X, Y) = s)$ pour les deux méthodes, Monte-Carlo et quasi-Monte-Carlo selon n , où $C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$, $a = -1.2$, $b = 0.6$ et $N_{MC} = N_{qMC} = 3000$ 68

RÉSUMÉ

Dans ce mémoire, nous explorons certaines façons d'effectuer le calcul de la vraisemblance d'une copule basée sur les rangs. Premièrement, nous exprimerons cette vraisemblance au moyen d'une espérance et d'une intégrale sur simplexe.

La valeur exacte de cette vraisemblance basée sur les rangs est difficile à obtenir de façon théorique. Pour avoir une approximation numérique de cette valeur, nous utiliserons donc deux algorithmes : Monte-Carlo et quasi-Monte-Carlo.

L'objectif est de comparer les deux algorithmes ; Monte-Carlo et quasi-Monte-Carlo à l'aide de la méthode de quadrature déterministe de Grundmann-Moller qui permet d'obtenir la valeur exacte de la vraisemblance basée sur les rangs des copules polynomiales. Parmi ces deux algorithmes, nous verrons lequel donne une meilleure approximation de cette vraisemblance selon d'autres paramètres comme la taille des rangs (la permutation) et la taille de l'échantillon de points utilisés dans ces deux méthodes.

Nous chercherons également à savoir jusqu'à quelle taille des rangs, l'ordinateur est capable de réaliser ce calcul via ces deux algorithmes.

MOTS-CLÉS : la vraisemblance d'une copule, basée sur les rangs, méthode de quasi-Monte-Carlo, méthode de Monte-Carlo, copule polynomiale, Grundmann-Moller.

INTRODUCTION

En statistique, la corrélation n'est qu'une mesure du lien linéaire entre les variables à l'étude, celle-ci ne repère pas les autres formes de liens. Pour récupérer la dépendance exacte, et ce, quel que soit le type de lien entre les variables, nous utilisons une fonction appelée copule. Tout récemment, Guillotte, Perron et Segers (2017) ont proposé un système innovateur de recommandation aux usagers (comme pour Netflix, Amazon, Ebay, etc.) à partir d'un modèle basé sur les copules. Ici, les variables représentent les notes (qui ne sont pas nécessairement observées) attribuées par les utilisateurs, et la copule exprime le lien entre les notes d'un expert et celles d'un usager en particulier. Ce qui nous intéresse ici, ce sont les rangs des objets, qui eux, sont observés, en entier pour l'expert et uniquement en partie pour l'usager. Le but est de recommander judicieusement un produit à l'usager, en inférant sur les rangs que l'on aurait observés si ce dernier avait classé tous les objets, au lieu d'une partie seulement.

Le point de départ est le calcul de la vraisemblance d'une copule basée sur les rangs, aux moyens d'algorithmes. Dans ce mémoire, nous explorons certaines méthodes d'effectuer ce calcul. Nous exprimerons tout d'abord cette vraisemblance au moyen d'une espérance, et donc, d'une intégrale sur un simplexe. Cette représentation suggère, d'une manière naturelle, certains algorithmes pour son évaluation numérique. Nous explorerons trois algorithmes : Grundmann-Moller, Monte-Carlo et quasi-Monte-Carlo.

Ce mémoire a pour objectif de comparer les deux méthodes, Monte-Carlo et quasi-Monte-Carlo, dans le calcul de la vraisemblance d'une copule basée sur les rangs,

tout en utilisant la méthode de Grundmann-Moller pour valider. Cette dernière donne la valeur exacte de la vraisemblance quand la copule est polynomiale.

Dans le premier chapitre, nous présenterons le concept de la copule et ses propriétés, ainsi que la simulation des observations distribuées selon une copule. Nous définirons ensuite, au chapitre II, quelques concepts spécifiques utilisés dans le calcul de cette vraisemblance comme la permutation, le simplexe, le rang, etc, et nous exprimerons cette vraisemblance au moyen d'une espérance et d'une intégrale sur un simplexe (Guillotte *et al.*, 2017). Le chapitre III portera sur la description et la théorie des trois méthodes (Monte-Carlo, quasi-Monte-Carlo et Grundmann-Moller) qui nous permettront de calculer ou approximer la vraisemblance.

Finalement, le quatrième chapitre de ce mémoire est consacré à l'objectif de ce travail, il traitera des comparaisons numériques des deux méthodes, Monte-Carlo et quasi-Monte-Carlo, en les validant par la méthode de Grundmann-Moller (qui donne la valeur exacte de la vraisemblance basée sur les rangs d'une copule polynomiale). Nous présenterons les résultats selon la taille des rangs (permutations), ainsi la taille de l'échantillon choisi pour les deux méthodes Monte-Carlo et quasi-Monte-Carlo. Nous explorerons aussi les limites de chaque méthode, plus précisément en étudiant jusqu'à quelle taille de rangs, ou quelle dimension, l'ordinateur est capable de calculer cette vraisemblance.

CHAPITRE I

LES COPULES

1.1 Introduction

Ce chapitre vise à présenter la notion de la copule, ainsi que la simulation qui sont parmi des concepts utilisés dans ce travail. Alors nous nous servirons de livre de Nelsen (2007), sous le titre «An introduction to copulas», pour écrire une grande partie de ce chapitre. Le lecteur pourra donc trouver plus de détails dans ce livre.

1.2 Fonction de répartition

Définition 1. Soit (X, Y) un vecteur aléatoire à valeurs dans R^2 .

La fonction de répartition bivariable F de (X, Y) est caractérisée par :

$$F(x, y) = P(X \leq x, Y \leq y), \text{ pour } (x, y) \in R^2.$$

Les fonctions des répartitions marginales s'obtiennent par la fonction de répartition conjointe comme suit :

$$\begin{aligned} F_X(x) &= \lim_{y \rightarrow +\infty} F(x, y) = F(x, +\infty), \\ \text{et } F_Y(y) &= \lim_{x \rightarrow +\infty} F(x, y) = F(+\infty, y). \end{aligned}$$

La fonction de densité conjointe f s'obtient de la fonction de répartition F comme

suit :

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}, \quad (x, y) \in \mathbb{R}^2.$$

Proposition 1.2.1. *Une fonction $F : \mathbb{R}^2 \rightarrow [0, 1]$ continue à droite est dite une fonction de répartition bivariée si :*

1. $\lim_{x \rightarrow -\infty} F(x, y) = 0$ et $\lim_{y \rightarrow -\infty} F(x, y) = 0$,
2. $\lim_{x_1, x_2 \rightarrow +\infty} F(x_1, x_2) = 1$,
3. pour tout $(a_1, a_2), (b_1, b_2) \in \mathbb{R}^2$, tel que $a_1 < b_1$ et $a_2 < b_2$, nous avons :

$$F(a_1, a_2) + F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) \geq 0.$$

1.3 Copules

Définition 2. *Nous appelons copule bivariée C , toute fonction de répartition bivariée définie sur $[0, 1]^2$, dont les lois marginales sont uniformes sur $[0, 1]$.*

Soit C est une copule bivariée, nous avons :

1. C est une fonction de répartition bivariée ;
2. $C(u, v) = 0$ si $u = 0$ ou $v = 0$;
3. $C(u, v) = u$ si $v = 1$, et $C(u, v) = v$ si $u = 1$.

Remarque 1.3.1. *Dans tout le reste du mémoire, nous utiliserons le mot «copule» au lieu de la « copule bivariée » pour alléger le texte.*

Théorème 1.3.2. *Soit C une copule. Pour tout $(u, v) \in [0, 1]^2$, nous avons :*

$$\max(0, u + v - 1) \leq C(u, v) \leq \min(u, v),$$

tel que $C_{\min}(u, v) = \max(0, u + v - 1)$ et $C_{\max}(u, v) = \min(u, v)$ sont des copules extrémales, appelées bornes de Fréchet.

Preuve :

Nous savons que : $P(U \leq u \text{ ou } V \leq v) \leq 1$,

donc : $P(U \leq u) + P(U \leq u) - P(U \leq u, V \leq v) \leq 1$,

d'où $u + v - 1 \leq C(u, v)$. Or $0 \leq C(u, v)$, donc $\max(0, u + v - 1) \leq C(u, v)$.

On a : $C(u, v) \leq C(u, 1) = u$ et $C(u, v) \leq C(1, v) = v$, donc $C(u, v) \leq \min(u, v)$.

1.4 Théorème de Sklar

Une autre façon de construire une copule consiste à utiliser la fonction de répartition bivariée. Le théorème de Sklar relie la copule à la fonction de répartition bivariée et aux lois marginales.

Théorème 1.4.1. (*Sklar, 1959*)

Si C est une copule, F_1 et F_2 sont deux fonctions de répartition, alors

$F(x, y) = C(F_1(x), F_2(y))$ est une fonction de répartition bivariée ayant F_1 et F_2 pour lois marginales.

Inversement, si F est une fonction de répartition conjointe des marges F_1 et F_2 , alors il existe une copule $C : [0, 1]^2 \rightarrow [0, 1]$, tel que pour tout $(x, y) \in \mathbb{R}^2$, nous avons :

$$F(x, y) = C(F_1(x), F_2(y)).$$

De plus, si F_1 et F_2 sont continues, alors C est unique.

Remarque 1.4.2. *Le théorème de Sklar permet d'exprimer une fonction de répartition bivariée F à l'aide d'une copule et de ses deux lois marginales F_1 et F_2 . À l'inverse, tel que nous le verrons dans le corollaire suivant, il est également possible d'exprimer une copule à partir de la fonction de répartition bivariée F et des inverses des fonctions de répartition marginales F_1 et F_2 .*

Corollaire 1.4.3. *Soit F une fonction de répartition conjointe de marges continues F_1 et F_2 , et C la copule associée à F . Alors, pour $(u, v) \in [0, 1]^2$, nous*

avons :

$$C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v)),$$

tel que $F_1^{-1}(u) = \inf\{x | F(x) \geq u\}$ et $F_2^{-1}(v) = \inf\{y | F(y) \geq v\}$.

À partir du corollaire 1.4.3, nous avons une formule de départ pour construire une copule à partir d'une fonction de répartition conjointe et de lois marginales.

Exemple

Soit la distribution logistique bivariee de Gumbel, dont la fonction de répartition conjointe est donnée par : $F(x, y) = (1 + e^{-x} + e^{-y})^{-1}$, pour tout $(x, y) \in \mathbb{R}^2$.

Nous obtenons les fonctions de répartition marginales :

$$F_1(x) = \lim_{y \rightarrow +\infty} F(x, y) = (1 + e^{-x})^{-1} \text{ et } F_2(y) = \lim_{x \rightarrow +\infty} F(x, y) = (1 + e^{-y})^{-1},$$

et les fonctions inverses sont respectivement :

$$F_1^{-1}(u) = -\ln(\frac{1}{u} - 1) \text{ et } F_2^{-1}(v) = -\ln(\frac{1}{v} - 1), \text{ pour tout } (u, v) \in [0, 1]^2.$$

En se basant sur le théorème de Sklar ; la copule associée à $F(x, y)$ est :

$$\begin{aligned} C(u, v) &= F(F_1^{-1}(u), F_2^{-1}(v)), \\ &= [1 + e^{\ln(\frac{1}{u}-1)} + e^{\ln(\frac{1}{v}-1)}]^{-1}, \\ &= \frac{uv}{u + v - uv}. \end{aligned}$$

1.5 Densité d'une copule

Si les fonctions de répartition marginales F_1 , F_2 et la copule C possèdent les densités respectives f_1 , f_2 et c , alors la densité conjointe f associée à la fonction de répartition F prend la forme suivante :

$$f(x, y) = f_1(x)f_2(y)c(F_1(x), F_2(y)), \quad (1.1)$$

où $f_1(x) = \frac{d}{dx}F_1(x)$, $f_2(x) = \frac{d}{dx}F_2(x)$ et $c(u, v) = \frac{\partial^2}{\partial u \partial v}C(u, v)$.

La formule (1.1) découle de celle du théorème de Sklar : $F(x, y) = C(F_1(x), F_2(y))$.

1.6 Exemples de copules

1.6.1 Copule d'indépendance (copule produit)

Si X et Y sont deux variables aléatoires indépendantes, nous avons :

$F(x, y) = F_1(x) F_2(y)$. Ainsi, le théorème suivant illustrera la copule d'indépendance.

Théorème 1.6.1. *Soient X_1 et X_2 deux variables aléatoires continues.*

X_1 et X_2 sont indépendantes si et seulement si la copule $C_{X_1, X_2}(u, v) = uv$.

La démonstration de ce théorème découle de celui de Sklar et du fait que

$$F(x, y) = F_1(x)F_2(y).$$

La densité de la copule d'indépendance : $c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v) = 1$.

1.6.2 Copule gaussienne

Une copule C est dite gaussienne de paramètre $\rho \in]-1, 1[$, si pour tout

$(u, v) \in [0, 1]^2$, nous avons :

$$C(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)),$$

tel que :

- Φ : fonction de répartition de la loi normale $N(0, 1)$;
- Φ_ρ : fonction de répartition du vecteur gaussien (X, Y) centrée de matrice de covariance $\begin{pmatrix} \rho & 1 \\ 1 & \rho \end{pmatrix}$ et $\Phi_\rho(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^x \int_{-\infty}^y \exp(-\frac{s^2+t^2-2\rho st}{2(1-\rho)}) ds dt$.

Nous pouvons, donc écrire la copule gaussienne sous forme :

$$C(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left(-\frac{s^2 + t^2 - 2\rho st}{2(1-\rho)}\right) ds dt.$$

La copule gaussienne est souvent caractérisée par sa densité c_ρ .

Soient ϕ et ϕ_ρ , respectivement la densité de loi normale $N(0, 1)$ et la densité gaussienne centrée de corrélation ρ . De la formule (1.1), nous déduisons que :

$$\phi_\rho(x_1, x_2) = \phi(x_1) \phi(x_2) c_\rho(\Phi(x_1), \Phi(x_2)).$$

Cela veut dire que :

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho)}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \times c_\rho(\Phi(x), \Phi(y)).$$

Donc, nous isolons $c_\rho(\Phi(x), \Phi(y))$, et après simplifications, nous avons :

$$c_\rho(\Phi(x), \Phi(y)) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho)} + \frac{x^2 + y^2}{2}\right),$$

ainsi, pour tout $(u, v) \in [0, 1]^2$, la densité de la copule gaussienne est sous la forme :

$$c_\rho(u, v) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{(\Phi^{-1}(u))^2 + (\Phi^{-1}(v))^2 - 2\rho(\Phi^{-1}(u))(\Phi^{-1}(v)))}{2(1-\rho)} + \frac{(\Phi^{-1}(u))^2 + (\Phi^{-1}(v))^2}{2}\right).$$

1.6.3 Copule archimédienne

Une copule C est dite archimédienne si elle est définie par :

$$C(u, v) = \phi^{-1}(\phi(u) + \phi(v)),$$

tel que ϕ est une fonction strictement décroissante et convexe de $[0, 1]$ dans \mathbb{R}^+ .

1.6.3.1 Exemples des copules archimédiennes

– Copule de Clayton

La copule de Clayton est définie par la fonction bivariee :

$$C_{\theta}(u, v) = \max(0, (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}),$$

où $\phi(u) = \frac{u^{-\theta}-1}{\theta}$ et le paramètre de la copule $\theta \in [-1, +\infty[\setminus\{0\}$.

– Copule de Gumbel

La copule de Gumbel est définie par la fonction bivariee :

$$C_{\theta}(u, v) = \exp[-((-\ln u)^{\theta} + (-\ln v)^{\theta})^{1/\theta}],$$

où $\phi(u) = (-\ln u)^{\theta}$ et le paramètre de la copule $\theta \in [1, +\infty[$.

– Copule de Frank

La copule de Frank est définie par la fonction bivariee :

$$C_{\theta}(u, v) = -\frac{1}{\theta} \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right],$$

où $\phi(u) = -\ln\left(\frac{e^{-\theta u}-1}{e^{-\theta}-1}\right)$ et $\theta \in \mathbb{R} \setminus \{0\}$.

1.6.4 Copules polynomiales

Nous présentons les deux types de copules polynomiales les plus connues, d'ordre deux et trois :

– Copule de Farlie-Gumbel-Morgenstern (Copule polynomiale d'ordre 2)

La copule de Farlie-Gumbel-Morgenstern est définie par :

$$C_{\theta}(u, v) = uv(1 + \theta(1 - u)(1 - v)),$$

où le paramètre $\theta \in [-1, 1]$ et sa copule densité associée est :

$$c_\theta(u, v) = 1 + \theta(1 - 2u)(1 - 2v).$$

– Famille de copules asymétriques (copule polynomiale d'ordre 3).

La famille des copules asymétriques avec des sections cubiques est définie par :

$$C(u, v) = uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b],$$

tel que $(a, b) \in \mathbb{R}^2$, $a \neq b$, $|b| \leq 1$ et $[b - 3 - (9 + 6b - 3b^2)^{1/2}]/2 \leq a \leq 1$.

1.7 Simulation

1.7.1 Génération d'observations (u, v) du vecteur aléatoire (U, V) d'une copule C

Pour générer des couples d'une copule C , nous remarquons que la fonction de répartition conditionnelle de V étant donné $U = u$ est donnée par :

$$\begin{aligned} P(V \leq v | U = u) &= \lim_{h \rightarrow 0} P(V \leq v | u \leq U \leq u + h), \\ &= \lim_{h \rightarrow 0} \frac{P(u \leq U \leq u + h, V \leq v)}{P(u \leq U \leq u + h)}, \\ &= \lim_{h \rightarrow 0} \frac{P(u \leq U \leq u + h, V \leq v)}{h}, \\ &\quad (\text{car la loi marginale d'une copule est une uniforme sur } [0, 1]), \\ &= \frac{\partial C(u, v)}{\partial u}, \\ &= C_u(v). \end{aligned}$$

Or la loi conjointe de (U, V) est le produit de la loi conditionnelle de V étant donné U et la loi marginale de U , nous obtenons :

Algorithme d'inversion,

- Générer (u, v') indépendantes, uniformes sur $(0, 1)$,

- Poser $v = C_u^{-1}(v')$ et obtenir (u, v) .

Afin de voir que V a la bonne loi marginale uniforme, il suffit d'intégrer :

$$\begin{aligned}
 P(V \leq v) &= \int_0^1 P(V \leq v | U = u) du, \\
 &= \int_0^1 \frac{\partial C(u, v)}{\partial u} du, \\
 &= C(1, v) - C(0, v), \\
 &= v - 0, \\
 &= v.
 \end{aligned}$$

1.7.2 Génération d'observations (u, v) du vecteur aléatoire (U, V) dont la copule est de Clayton

La copule de Clayton est définie par : $C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$ où $\theta \in]0, +\infty[$. Nous avons :

$$\begin{aligned}
 C_u(v) &= \frac{\partial C_\theta(u, v)}{\partial u}, \\
 &= u^{-\theta-1} (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}-1}.
 \end{aligned}$$

L'inverse est : $C_u^{-1}(v') = [(v' u^{\theta+1})^{-\frac{\theta}{1+\theta}} + 1 - u^{-\theta}]^{-1/\theta}$.

Pour générer des couples (u, v) de la copule Clayton :

- Nous générons des couples (u, v') indépendantes, uniformes sur $(0, 1)^2$,
- Nous posons $v = C_u^{-1}(v') = [(v' u^{\theta+1})^{-\frac{\theta}{1+\theta}} + 1 - u^{-\theta}]^{-1/\theta}$, et nous obtenons les couples (u, v) .

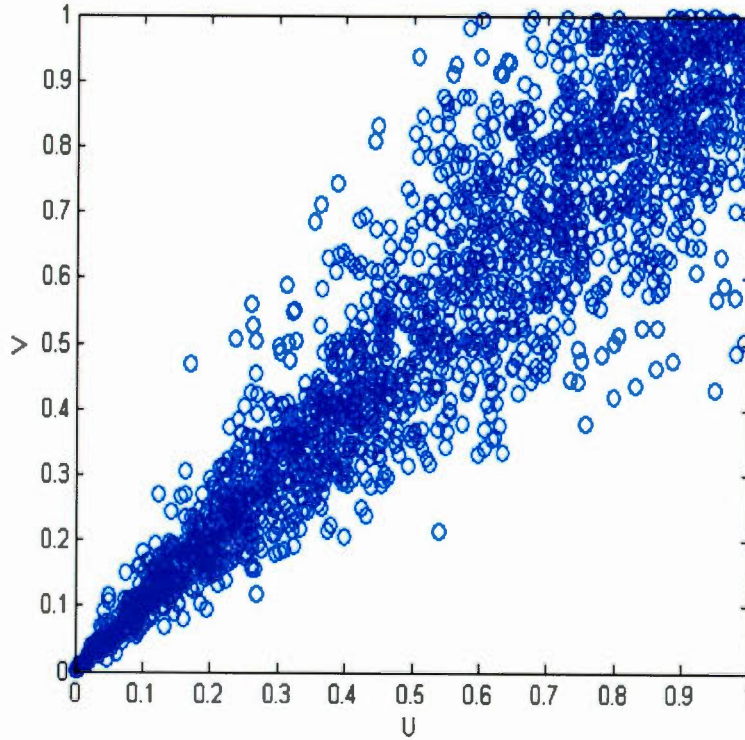


Figure 1.1 Génération d'observations (u, v) de la copule Clayton, pour $\theta = 7$.

1.7.3 Génération d'observations (x_1, x_2) du vecteur aléatoire (X_1, X_2) dont la fonction de répartition F a les marges continues

Soit X_1 et X_2 deux variables aléatoires de fonction de répartition conjointe F . Soit F_1 et F_2 leurs fonctions de répartition marginales continues respectivement, et C leur copule. Ces deux relations :

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \quad \text{et} \quad C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v)),$$

nous illustrent comment générer les couples (x_1, x_2) de la fonction de répartition F .

Algorithme d'inversion :

- Nous générons (u, v) selon C ; comme nous avons décrit dans la section (1.6.1).
- Nous posons $x_1 = F_1^{-1}(u)$ et $x_2 = F_2^{-1}(v)$ et nous obtenons les couples (x_1, x_2) .

Les variables X_1 et X_2 ont les bonnes lois marginales, de copule identique (par invariance : F_1^{-1} et F_2^{-1} sont strictement croissantes ici) à (U, V) qui est C .

CHAPITRE II

LA VRAISEMBLANCE BASÉE SUR LES RANGS

2.1 Introduction

Nous définirons dans ce chapitre quelques concepts spécifiques (la permutation, le simplexe, les rangs, etc) qui sont directement liés à notre objectif qui est le calcul de la vraisemblance, et nous exprimerons cette vraisemblance au moyen d'une espérance et d'une intégrale sur un simplexe. Nous nous servons de l'article, qui est en cours de publication, de Guillotte *et al.* (2017) pour écrire la grande partie de ce chapitre.

2.2 Permutation

Définition 3. *Une permutation σ d'un ensemble A , de cardinal fini, est l'une des bijections de A à lui-même.*

Soit S_n l'ensemble des permutations de $\{1, \dots, n\}$, muni d'une loi de composition interne \circ ;

- un élément $\sigma \in S_n$ est noté par $\sigma = (\sigma(1), \dots, \sigma(n))$,
- soit $\sigma_1, \sigma_2 \in S_n$. $\sigma_1 \circ \sigma_2 \in S_n$ est défini par $(\sigma_1 \circ \sigma_2)(i) = \sigma_1(\sigma_2(i))$,
pour $i = 1, \dots, n$,
- l'élément neutre de S_n est : $e = (1, 2, \dots, n)$, nous avons : $\sigma \circ e = e \circ \sigma = \sigma$,

- l'inverse de $\sigma \in S_n$ est noté σ^{-1} , tel que : $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = e$.

La représentation matricielle de la permutation $\sigma \in S_n$ est :

$$M_\sigma = \begin{pmatrix} e_{\sigma(1)}^\top \\ \vdots \\ e_{\sigma(n)}^\top \end{pmatrix},$$

où $e_i, i = 1, \dots, n$, est le i^e élément de la base standard de \mathbb{R}^n .

Nous avons aussi :

$$M_{\sigma_1 \circ \sigma_2} = M_{\sigma_1} M_{\sigma_2}, \quad M_{\sigma^{-1}} = (M_\sigma)^{-1} \quad \text{et} \quad M_e = I_n,$$

où $e = (1, \dots, n)$ est la permutation neutre de S_n muni de la loi de composition \circ , et I_n est la matrice identité d'ordre n .

Soit $x \in \mathbb{R}^n$ et $x = (x_1, x_2, \dots, x_n)^\top$, nous avons : $M_\sigma x = (x_{\sigma(1)}, \dots, x_{\sigma(n)})^\top$.

Exemple

Si $n = 4$, nous avons : $e_3 = (0, 0, 1, 0)^\top$.

Soit $\sigma = \begin{pmatrix} 3 & 1 & 4 & 2 \end{pmatrix}$. Alors :

$$\sigma^{-1} = \begin{pmatrix} 2 & 4 & 1 & 3 \end{pmatrix}, \quad \sigma \circ \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix},$$

$$M_\sigma = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{et} \quad M_{\sigma^{-1}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = (M_\sigma)^{-1}.$$

Si $x = \begin{pmatrix} 10 & 23 & 1 & 4 \end{pmatrix}^\top$, alors :

$$M_\sigma x = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 10 \\ 23 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \\ 4 \\ 23 \end{pmatrix}.$$

2.3 Statistiques du rang

Soit $x = (x_1, x_2, \dots, x_n)^\top$ un vecteur aléatoire de \mathbb{R}^n .

Soit $D_n = \{x \in \mathbb{R}^n : x_{(1)} < \dots < x_{(n)}\}$ un ensemble de vecteurs qui ont des composantes deux à deux distinctes, où $x_{(i)}$ est la i^e statistique d'ordre de (x_1, \dots, x_n) , $i = 1, \dots, n$.

Soit $x \in D_n$, la statistique du rang de x est définie par :

$$r_x = \text{rang}(x) = (\text{rang}(x_1), \dots, \text{rang}(x_n)),$$

où $\text{rang}(x_i)$ est le rang de x_i parmi x_1, \dots, x_n .

Pour tout $x \in D_n$, nous pouvons considérer la statistique du rang r_x comme une permutation dans S_n , vu que les éléments de S_n sont sous la forme :

$$\sigma = (\sigma(1), \dots, \sigma(n)).$$

Les statistiques ci-dessous sont utiles pour ce qui suit :

$$R(x, y) = \begin{cases} (r_x, r_y)^\top & \text{if } (x, y) \in D_n^2, \\ (e, e)^\top & \text{sinon,} \end{cases} \quad \text{et} \quad S(x, y) = \begin{cases} r_y \circ r_x^{-1} & \text{si } (x, y) \in D_n^2, \\ e & \text{sinon.} \end{cases} \quad (2.1)$$

Notons que nous avons :

$$\{(x, y) \in \mathbb{R}^{2n} : S(x, y) = s\} = \bigcup_{\sigma \in S_n} \{(x, y) \in \mathbb{R}^{2n} : R(x, y) = (\sigma, s \circ \sigma)^\top\}, \quad s \in S_n.$$

Remarque 2.3.1. $S(x, y)$ représente les rangs des y présentés dans l'ordre des valeurs des x dans D_n .

Exemple :

Soit : $x = (40, 10, -2, 30)$ et $y = (45, 22, 100, 60)$.

Nous avons :

$$R(x, y) = \begin{pmatrix} r_x \\ r_y \end{pmatrix} = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 2 & 1 & 4 & 3 \end{pmatrix} \in S_4^2,$$

$$\text{et } r_x^{-1} = \begin{pmatrix} 4 & 2 & 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & 2 & 4 & 1 \end{pmatrix}.$$

Donc :

$$S(x, y) = r_y \circ r_x^{-1} = \begin{pmatrix} 2 & 1 & 4 & 3 \end{pmatrix} \circ \begin{pmatrix} 3 & 2 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 3 & 2 \end{pmatrix}.$$

2.4 Énoncé du problème

Nous avons un échantillon de variables indépendantes et identiquement distribuées, $(X_1, Y_1), \dots, (X_n, Y_n)$, dont la fonction de répartition conjointe a C pour copule.

Nous cherchons alors à calculer la probabilité $P(R_X = r_x, R_Y = r_y)$.

$$\begin{aligned} P(R_X = r_x, R_Y = r_y) &= P(R(X, Y) = R(x, y)), \\ &= \frac{1}{n!} P(S(X, Y) = S(x, y)), \\ &= \frac{1}{n!} P(S(X, Y) = s). \end{aligned}$$

2.5 La vraisemblance des rangs

2.5.1 Exprimer la vraisemblance basée sur les rangs par une espérance

Soit \mathcal{H} un espace des fonctions de répartition bivariées tel que ses marginaux soient continus.

Supposons que (X_i, Y_i) , $i = 1, \dots, n$, sont des bivariées indépendantes et identiquement distribuées et leur fonction de répartition conjointe $H \in \mathcal{H}$,

$$H(x, y) = P(X_1 \leq x, Y_1 \leq y), \text{ pour } (x, y) \in \mathbb{R}^2.$$

$F(x) = H(x, \infty)$ pour $x \in \mathbb{R}$ et $G(y) = H(\infty, y)$ pour $y \in \mathbb{R}$, des fonctions marginales continues pour chaque variable.

Notons que dans ce cas, nous avons : $P(X_i = X_j) = 0$ et $P(Y_i = Y_j) = 0$, pour tout i et j tel que $i \neq j$.

Si nous posons $U_1 = F(X_1)$ et $V_1 = G(Y_1)$, alors :

$$C(u, v) = P(U_1 \leq u, V_1 \leq v), \text{ pour } (u, v) \in [0, 1]^2,$$

est une copule (une fonction de répartition bivariable avec des marginaux uniformes).

Donc d'après le théorème de Sklar, H admet la représentation :

$$H(x, y) = C(F(x), G(y)), (x, y) \in \mathbb{R}^2. \quad (2.2)$$

Proposition 2.5.1. (*Guillotte et al., 2017*)

Soient (X_i, Y_i) , $i = 1, \dots, n$, des bivariées indépendantes et identiquement distribuées ayant une fonction de répartition $H \in \mathcal{H}$, avec la copule C , et leurs fonctions de répartition marginales (continues) respectivement F et G .

Si la copule C a une fonction de densité c , nous avons donc :

$$P(S(X, Y) = s) = \frac{1}{n!} \mathbb{E} \left(\prod_{i=1}^n c(U_{(i)}, V_{(s(i))}) \right), \quad (2.3)$$

où $U_{(1)}, \dots, U_{(n)}$ et $V_{(1)}, \dots, V_{(n)}$ sont respectivement des statistiques d'ordre de U_1, \dots, U_n et V_1, \dots, V_n , deux échantillons de variables aléatoires indépendantes de la distribution uniforme sur $(0, 1)$.

La proposition 2.5.1 est démontrée à l'appendice A.

Le résultat (2.3) de la proposition 2.5.1 démontre spécifiquement que les distributions marginales de deux variables X et Y n'affectent pas la probabilité des rangs,

comme nous aurions pu le prévoir. Nous supposons donc qu'elles sont tous les deux distribuées uniformément sur $(0, 1)^n$. (Guillotte *et al.*, 2017).

Lorsque la fonction de la copule appartient à une famille paramétrique : $\{C_\theta : \theta \in \Theta\}$, alors cette probabilité, qui est en fonction de θ , est parfois appelée la vraisemblance des rangs.

L'objectif de ce mémoire est justement de trouver un moyen de calculer une approximation de la vraisemblance $P(S(X, Y) = s)$.

À partir de l'expression (2.3) et d'après la loi forte des grands nombres, nous pourrions remarquer qu'il y a une manière d'approximer la valeur de cette vraisemblance en utilisant deux méthodes statistiques, telles que Monte-Carlo et quasi-Monte-Carlo que nous aborderons avec plus de détails dans le prochain chapitre.

Dans ce qui suit, nous illustrerons une autre expression de la vraisemblance, ainsi qu'un autre moyen de calculer une approximation de cette vraisemblance.

2.5.2 Exprimer la vraisemblance par une intégrale sur un simplexe

Notons Δ_n un simplexe standard de dimension n donné par :

$$\Delta_n = \left\{ (x_1, \dots, x_n) \in [0, 1]^n : \sum_{i=1}^n x_i \leq 1 \right\}.$$

Pour une permutation $s \in S_n$, nous notons aussi :

$$E_s = \{(u, v) \in [0, 1]^{2n} : R(u, v) = (e, s)^\top\}.$$

Proposition 2.5.2. (Guillotte *et al.*, 2017)

Soient (X_i, Y_i) , $i = 1, \dots, n$, des bivariées indépendantes et identiquement distribuées ayant une fonction de répartition $H \in \mathcal{H}$ avec la copule C , et leurs fonctions de répartition marginales (continues) sont respectivement F et G .

Si la copule C a une fonction de densité c , nous avons donc :

$$P(S(X, Y) = s) = n! \int_{\Delta_n^2} c(\varphi_s(x, y)) dx dy, \quad (2.4)$$

où φ_s est une bijection définie par $\varphi_s : \Delta_n \times \Delta_n \rightarrow E_s$,

$$\varphi_s(x, y) = (Lx, M_s Ly),$$

tel que L est une matrice triangulaire inférieure :

$$L = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 1 & 1 & \ddots & & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & & & 0 \\ 1 & 1 & \cdots & \cdots & 1 \end{pmatrix},$$

et M_s est une matrice de permutation qui correspond à $s \in S_n$. Le Jacobien de φ_s est égal à 1.

La proposition 2.5.2 est démontrée à l'appendice B.

L'expression (2.4) nous suggère d'exploiter une formule d'intégration invariante sur le simplexe, qui a été faite par Grundmann et Möller (1978), pour approximer la valeur de cette vraisemblance ou pour évaluer la valeur exacte de la vraisemblance d'une copule polynomiale. Nous détaillerons cette méthode dans le prochain chapitre.

CHAPITRE III

MÉTHODES UTILISÉES POUR APPROXIMER LA VRAISEMBLANCE BASÉE SUR LES RANG

3.1 Méthode de Grundmann-Möller

Selon l'article publié en 1978 par Grundmann et Möller, intitulé «Invariant integration formulas for the n -simplex by combinatorial methods», la méthode de Grundmann-Möller est une méthode de quadrature qui permet de calculer une approximation de l'intégrale d'une fonction sur un simplexe. La particularité de cette méthode c'est qu'elle rend possible l'évaluation d'une intégrale d'un monôme sur un simplexe de manière exacte et sans aucune erreur, raison pour laquelle nous l'utiliserons pour calculer la valeur exacte de la vraisemblance d'une copule polynomiale basée sur les rangs. Alors nous nous servons de l'article de Grundmann et Möller (1978) pour écrire cette section.

3.1.1 L'intégrale d'un monôme sur un simplexe de dimension n

Soit x une variable de \mathbb{R}^n . Nous considérons l'intégrale d'une fonction f sur un simplexe de dimension n :

$$\begin{aligned}
I_n f &= \int_{\Delta_n} f(x) dx, \\
&= \int_0^1 \int_0^{1-x_1} \cdots \int_0^{1-x_1-\cdots-x_{n-1}} f(x_1, \dots, x_n) dx_n dx_{n-1} \cdots dx_1.
\end{aligned}$$

Proposition 3.1.1. (*Stroud, 1971*)

Pour tout $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, le monôme

$$\begin{aligned}
x^\alpha : \quad \mathbb{R}^n &\rightarrow \mathbb{R} \\
(x_1, \dots, x_n) &\rightarrow x_1^{\alpha_1} \cdots x_n^{\alpha_n}
\end{aligned}$$

satisfait l'équation suivante :

$$I_n x^\alpha = \frac{\alpha_1! \cdots \alpha_n!}{(n + |\alpha|)!}, \quad (3.1)$$

avec $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

Si on définit la fonction X^α , pour tout $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, par :

$$\begin{aligned}
X^\alpha : \quad \mathbb{R}^n &\rightarrow \mathbb{R} \\
(x_1, \dots, x_n) &\rightarrow \left(1 - \sum_{i=1}^n x_i\right)^{\alpha_0} x_1^{\alpha_1} \cdots x_n^{\alpha_n},
\end{aligned}$$

nous obtiendrons alors un résultat similaire à (3.1) :

$$I_n X^\alpha = \frac{\alpha_0! \alpha_1! \cdots \alpha_n!}{(n + |\alpha|)!}, \quad (3.2)$$

avec $|\alpha| = \alpha_0 + \alpha_1 + \cdots + \alpha_n$.

Exemple

Si $n = 3$ et $\alpha = (2, 4, 1)$, l'intégrale de monôme $x^\alpha = x_1^2 x_2^4 x_3^1$ sur un simplexe de dimension 3 est :

$$\begin{aligned}
 I_3 x^\alpha &= \int_{\Delta_3} x^\alpha dx \\
 &= \int_{\Delta_3} x_1^2 x_2^4 x_3^1 dx_1 dx_2 dx_3 \\
 &= \frac{2! 4! 1!}{(3 + (2 + 4 + 1))} \quad (\text{d'après (3.1)}) \\
 &= \frac{24}{5}.
 \end{aligned}$$

Proposition 3.1.2. (Grundmann et Möller, 1978)

Soit $d \in \mathbb{N}$ et $\alpha \in \mathbb{N}^n$. L'espace linéaire \mathcal{P}_d de polynômes de degré au plus d , qui est engendré par les monômes x^α tel que $|\alpha| \leq d$, admet $\{X^\alpha / |\alpha| = d\}$ pour base. Cette dernière est invariante sous toutes les transformations affines du simplexe

$$\Delta_n = \{(x_1, \dots, x_n) \in [0, 1]^n : \sum_{i=1}^n x_i \leq 1\}$$

dans lui-même, C'est-à-dire invariant sous les transformations :

$$\begin{aligned}
 \varphi_\sigma : \quad \Delta_n &\rightarrow \Delta_n \\
 (x_1, \dots, x_n) &\rightarrow (x_{\sigma(1)}, \dots, x_{\sigma(n)}),
 \end{aligned}$$

avec $\sigma = (\sigma(0), \dots, \sigma(n))$ une permutation de $\{0, \dots, n\}$ et $x_0 = 1 - \sum_{i=1}^n x_i$.

Voir la preuve dans l'article de Grundmann et Möller (1978).

Exemple

Soit $n = 3$, la permutation $\sigma = (2, 0, 3, 1)$ et $x = (0.25, 0.1, 0.35)$. Nous avons $x \in \Delta_3$, car $0.25 + 0.1 + 0.35 \leq 1$.

On a : $x_1 = 0.25$, $x_2 = 0.1$, $x_3 = 0.35$, et $x_0 = 1 - (0.25 + 0.1 + 0.35) = 0.3$.

Donc : $M_\sigma(0.3, 0.25, 0.1, 0.35)^T = (0.25, 0.35, 0.3, 0.1)$.

Et par suite : $\varphi_\sigma(x) = \varphi_\sigma(0.25, 0.1, 0.35) = (0.35, 0.3, 0.1)$.

Définition 4. Pour tout point $x = (x_1, \dots, x_n) \in \Delta_n$, on définit une classe d'équivalence de x , composée de tous les points qui sont l'image de x par les transformations φ_σ . On la note par $((1 - \sum_1^n x_i, x_1, \dots, x_n))$ ou brièvement par $((x))$. Pour une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, on note :

$$\sum f((x)) = \sum_{y \in ((x))} f(y),$$

où y est un élément de la classe d'équivalence de x .

Exemple

Soit $x = (0.5, 0.2)$. La classe d'équivalence de x est :

$$((x)) = (((0.5, 0.2))) = \{(0.5, 0.2), (0.2, 0.5), (0.5, 0.3), (0.3, 0.5), (0.2, 0.3), (0.3, 0.2)\}.$$

Ainsi, si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, nous avons :

$$\sum f((x)) = f(0.5, 0.2) + f(0.2, 0.5) + f(0.5, 0.3) + f(0.3, 0.5) + f(0.2, 0.3) + f(0.3, 0.2).$$

3.1.2 La formule d'intégration pour un simplexe de dimension n

Définition 5. (Grundmann et Möller, 1978)

Soit une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Pour $A_1, \dots, A_m \in \mathbb{R}$ et les classes d'équivalences $((x^{(1)})), \dots, ((x^{(m)}))$ qui sont différentes. On dit que

$$\sum_{i=1}^m A_i \sum f((x^{(i)})) = I_n f - Rf$$

est une formule d'intégration invariante de degrés d si $Rf = 0$ pour tout $f \in \mathcal{P}_d$,

où \mathcal{P}_d est l'espace linéaire de polynômes de degrés au plus d .

Si c'est le cas, A_1, \dots, A_m sont les poids et $((x^{(1)})), \dots, ((x^{(m)}))$ sont les nœuds de la formule d'intégration invariante.

Théorème 3.1.3. (Grundmann et Möller, 1978)

Soit une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $r \in \mathbb{N}$, $d = 2r + 1$ et $n \in \mathbb{N}$.

Donc :

$$\sum_{i=0}^r (-1)^i 2^{-2r} \frac{(d+n-2i)^d}{i!(d+n-i)!} \sum_{\substack{\beta_0 \geq \dots \geq \beta_n \\ |\beta|=r-i}} f\left(\frac{2\beta_0+1}{d+n-2i}, \dots, \frac{2\beta_n+1}{d+n-2i}\right) = I_n f - R_{nd} f \quad (3.3)$$

est une formule d'intégration invariante de degrés d .

Où $\beta = (\beta_0, \dots, \beta_n) \in \mathbb{N}^{n+1}$ est l'indice de la deuxième somme de la formule, $|\beta| = |\beta_0 + \dots + \beta_n|$, $I_n f = \int_{\Delta_n} f(x) dx$, $x \in \mathbb{R}^n$ et $R_{nd} f$ est l'erreur de la formule d'intégration.

Lemme 3.1.4. On trouve $R_{nd} X^\alpha = 0$ pour tout $\alpha \in \mathbb{N}^{n+1}$, tel que $|\alpha| = d$.

Cela veut dire que $R_{nd} x^\eta = 0$ pour tout $\eta \leq d$, car

$x^\eta \in$ generateur $\{X^\alpha / |\alpha| = d\} = \mathcal{P}_d$ pour tout $\eta \in \mathbb{N}^n$, tel que $|\eta| \leq d$.

Remarque 3.1.5. Le nombre de nœuds pour cette formule d'intégration est :

$$N_r = \sum_{i=0}^r \binom{n+r-i}{r} = \binom{n+r+1}{r}.$$

Exemples

Soit une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Nous déterminons la formule d'intégration invariante de degrés d pour la fonction f en utilisant la formule (3.3);

– Si nous prenons $r = 0$, donc $d = 2 \times 0 + 1 = 1$ et la formule d'intégration invariante de f de degré $d = 1$ est :

$$\frac{1}{n!} \sum f\left(\frac{1}{n+1}, \dots, \frac{1}{n+1}\right) = \frac{1}{n!} f\left(\frac{1}{n+1}, \dots, \frac{1}{n+1}\right) = I_n f - R_{n1} f.$$

Nous avons un seul nœud $(\frac{1}{n+1}, \dots, \frac{1}{n+1})$ et un seul poids $\frac{1}{n!}$.

– Si nous prenons $r = 1$, donc $d = 2 \times 1 + 1 = 3$ et la formule d'intégration invariante de f de degré $d = 3$ est :

$$\begin{aligned} \frac{1}{4} \frac{(n+3)^3}{(n+3)!} \sum f\left(\left(\frac{3}{n+3}, \frac{1}{n+3}, \dots, \frac{1}{n+3}\right)\right) - \frac{1}{4} \frac{(n+1)^3}{(n+2)!} \sum f\left(\left(\frac{1}{n+1}, \dots, \frac{1}{n+1}\right)\right) \\ = I_n f - R_{n3} f. \end{aligned}$$

Nous avons $\binom{n+1+1}{1} = n+2$ nœuds dans cette formule d'intégration.

Si nous posons $n = 2$, la formule d'intégration invariante de f de degré $d = 3$ sera :

$$\frac{1}{4} \frac{(2+3)^3}{(2+3)!} \sum f\left(\left(\frac{3}{2+3}, \frac{1}{2+3}\right)\right) - \frac{1}{4} \frac{(2+1)^3}{(2+2)!} \sum f\left(\left(\frac{1}{2+1}, \frac{1}{2+1}\right)\right) = I_2 f - R_{2,3} f,$$

$$\frac{25}{96} \sum f\left(\left(\frac{3}{5}, \frac{1}{5}\right)\right) - \frac{3}{32} \sum f\left(\left(\frac{1}{3}, \frac{1}{3}\right)\right) = I_2 f - R_{2,3} f,$$

$$\frac{25}{96} f\left(\frac{3}{5}, \frac{1}{5}\right) + \frac{25}{96} f\left(\frac{1}{5}, \frac{3}{5}\right) + \frac{25}{96} f\left(\frac{1}{5}, \frac{1}{5}\right) - \frac{3}{32} f\left(\frac{1}{3}, \frac{1}{3}\right) = I_2 f - R_{2,3} f.$$

Nous avons donc quatre nœuds au total. Trois nœuds $(\frac{3}{5}, \frac{1}{5})$, $(\frac{1}{5}, \frac{3}{5})$, $(\frac{1}{5}, \frac{1}{5})$ qui ont un poids en commun soit $\frac{25}{96}$, et un nœud $(\frac{1}{3}, \frac{1}{3})$ qui a le poids $-\frac{3}{32}$.

Dans ce cas, le degré de la formule d'intégration invariante est à $d = 3$. Alors d'après le théorème 3.1.4, si f appartient à \mathcal{P}_3 , nous aurons $R_{2,3} f = 0$, et

$$I_2 f = \int_{\Delta_2} f(x) dx = \frac{25}{96} f\left(\frac{3}{5}, \frac{1}{5}\right) + \frac{25}{96} f\left(\frac{1}{5}, \frac{3}{5}\right) + \frac{25}{96} f\left(\frac{1}{5}, \frac{1}{5}\right) - \frac{3}{32} f\left(\frac{1}{3}, \frac{1}{3}\right).$$

Par exemple :

$$\begin{aligned} \text{Soit } f : \mathbb{R}^2 &\rightarrow \mathbb{R}, \\ (x_1, x_2) &\rightarrow 3x_1^2x_2 + 2x_2^2 + 4x_1x_2 + 6. \end{aligned}$$

Puisque $f \in \mathcal{P}_3$, l'erreur de la formule de l'intégration invariante $R_{2,3}f = 0$, et

$$I_2f = \int_{\Delta_2} f(x_1, x_2) dx_1 dx_2 = \frac{25}{96} f\left(\frac{3}{5}, \frac{1}{5}\right) + \frac{25}{96} f\left(\frac{1}{5}, \frac{3}{5}\right) + \frac{25}{96} f\left(\frac{1}{5}, \frac{1}{5}\right) - \frac{3}{32} f\left(\frac{1}{3}, \frac{1}{3}\right).$$

En général, si $f \in \mathcal{P}_d$, tel que $d = 2r + 1$ et $r \in \mathbb{N}$ et $n \in \mathbb{N}$, nous pourrions évaluer $I_n f$ par le premier terme de la formule d'intégration invariante du théorème 3.1.4 :

$$I_n f = \sum_{i=0}^r (-1)^i 2^{-2r} \frac{(d+n-2i)^d}{i!(d+n-i)!} \sum_{|\beta|=r-i}^{\beta_0 \geq \dots \geq \beta_n} \sum f\left(\frac{2\beta_0+1}{d+n-2i}, \dots, \frac{2\beta_n+1}{d+n-2i}\right).$$

Si $f \notin \mathcal{P}_d$, nous pourrions seulement approximer $I_n f$ par le premier terme de la formule d'intégration invariante du théorème 3.1.4 :

$$I_n f \approx \sum_{i=0}^r (-1)^i 2^{-2r} \frac{(d+n-2i)^d}{i!(d+n-i)!} \sum_{|\beta|=r-i}^{\beta_0 \geq \dots \geq \beta_n} \sum f\left(\frac{2\beta_0+1}{d+n-2i}, \dots, \frac{2\beta_n+1}{d+n-2i}\right).$$

Quand le degré de la formule d'intégration invariante d augmente, le nombre de nœuds augmente, et la marge d'erreur $|R_{nd}f|$ diminue.

3.2 Méthode de Monte-Carlo

3.2.1 Introduction

Dans cette section, nous parlerons de la description et de la théorie de la méthode de Monte-Carlo en se servant en grande partie de la thèse de Thiémard (2000).

On désigne par le terme «méthode de Monte-Carlo» toute méthode ayant pour but de calculer une valeur numérique en se servant de procédés aléatoires, en d'autres

termes, en utilisant des techniques probabilistes. Il est nécessaire de souligner qu'il n'existe pas de définition précise de cette technique nommée Monte-Carlo, mais afin de la décrire, on peut dire que les méthodes de ce type se caractérisent essentiellement par l'utilisation du hasard afin de résoudre des problèmes centrés sur le calcul d'une valeur numérique. La réponse obtenue par cette méthode aura la forme d'une réponse statistique.

Le progrès informatique qui a considérablement augmenté la puissance des ordinateurs a rendu ces méthodes opérationnelles. Ainsi, la méthode de Monte-Carlo est préférentiellement utilisée dans de nombreux secteurs scientifiques et technologiques tel que les mathématiques, la finance, la physique, la génétique et bien d'autres encore. D'une manière générale, l'utilisation de cette méthode concerne tous les domaines où l'utilisation des méthodes scientifiques se heurte à des difficultés. Dans ce contexte, les problèmes déterministes, qui font appel aux calculs numériques, constituent la majeure partie des domaines où la méthode de Monte-Carlo peut être utilisée avec succès (exemples : calculs d'intégrales multiples, estimation des surfaces, résolution d'équations différentielles, ...). (Thiémard, 2000).

3.2.2 Principe de la méthode de Monte-Carlo

La première étape pour calculer une quantité via la méthode de Monte-Carlo est de la mettre sous forme d'une espérance, afin de passer à l'étape suivante qui consiste à calculer cette quantité par une espérance $\mathbb{E}(X)$ de la variable aléatoire X . Pour réaliser ce calcul, il convient de savoir simuler une variable aléatoire selon la loi de X .

On dispose donc d'une suite $(x_i)_{1 \leq i \leq N}$ de N réalisations de la variable aléatoire X . Alors d'après la loi des grands nombres, nous effectuons une approximation de $\mathbb{E}(X)$ par :

$$\mathbb{E}(X) \approx \frac{1}{N}(x_1 + \dots + x_N).$$

3.2.3 Application de la méthode de Monte-Carlo au calcul d'une intégrale sur $[0, 1]^n$

3.2.3.1 Calcul de l'intégrale par l'espérance

Soit $f : [0, 1]^n \rightarrow \mathbb{R}$, une fonction intégrable sur $[0, 1]^n$. Nous cherchons à calculer l'intégrale :

$$\begin{aligned} I &= \int_{[0,1]^n} f(x) dx, \\ &= \int_{[0,1]^n} f(x_1, \dots, x_n) dx_1 \cdots dx_n. \end{aligned}$$

Dans ce cas, la méthode de Monte-Carlo consiste à écrire cette intégrale sous forme d'espérance de f qui est la généralisation de la moyenne de f sur $[0, 1]^n$, et si $U = (U_1, \dots, U_n)$ est un vecteur de variables aléatoires uniformément distribuée sur $[0, 1]^n$, alors l'espérance de la variable aléatoire $f(U)$ est égale à la valeur de cette intégrale $I = \int_{[0,1]^n} f(x) dx$:

$$I = \mathbb{E}(f(U)) = \mathbb{E}(f(U_1, \dots, U_n)).$$

Cette étape de transformation consiste à contourner le problème du calcul d'une intégrale en le remplaçant par celui de la détermination d'une espérance. À l'issue de cette étape, nous disposons alors d'une suite $(u^i)_{1 \leq i \leq N}$ de N réalisations indépendantes du vecteur U uniformément distribuées sur $[0, 1]^n$, tel que

$$u^i = (u_1^i, \dots, u_n^i) \quad \text{pour } i = 1, \dots, N.$$

Nous approximations l'intégrale I comme suit :

$$I \approx I_N = \frac{1}{N} \sum_{i=1}^N f(u^i).$$

Quand N le nombre de points u^i augmente, l'approximation est plus précise et nous avons :

$$I = \int_{[0,1]^n} f(x)dx = \lim_{N \rightarrow \infty} I_N = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(u^i),$$

où $(u^i)_{1 \leq i \leq N}$ est un échantillon de points simulés selon la loi uniforme sur $[0, 1]^n$.

Ce résultat découle de la loi forte des grands nombres.

3.2.3.2 Convergence de l'approximation d'une intégrale par la méthode de Monte-Carlo

Justifier la convergence de la méthode de Monte-Carlo est rendu possible grâce à la loi forte des grands nombres. Le théorème de la limite centrale précise la vitesse de convergence.

Soit $(u^i)_{i \geq 1}$ une suite de réalisations du vecteur aléatoire U qui est uniformément distribué sur $[0, 1]^n$. Alors, d'après la loi des grands nombres, nous avons :

$$\lim_{N \rightarrow +\infty} \frac{1}{N} (f(u^1) + \dots + f(u^N)) = \mathbb{E}(f(U)) = I.$$

Pour évaluer l'intérêt de cette méthode, il est essentiel de pouvoir évaluer l'erreur commise, qui est définie par :

$$\begin{aligned} R_N &= \mathbb{E}(f(U)) - \frac{1}{N} (f(u^1) + \dots + f(u^N)), \\ &= I - I_N. \end{aligned}$$

De plus, si la variance de la fonction $f(U)$; $\sigma^2 = \int_{[0,1]^n} (f(x) - \mathbb{E}(f(U)))^2 dx$, est finie, nous obtenons :

$$Var(I_N) = \frac{1}{N^2} \sum_{i=1}^N Var(f(U^i)) = \frac{\sigma^2}{N},$$

comme variance de l'approximation de l'intégrale par la méthode de Monte-Carlo I_N .

Pour de grandes valeurs de N , le théorème de la limite central peut s'imposer afin de construire un intervalle de confiance de l'intégrale I du type :

$$\left[\frac{1}{N} \sum_{i=1}^N f(u^i) - z_\alpha \frac{\sigma}{\sqrt{N}}, \frac{1}{N} \sum_{i=1}^N f(u^i) + z_\alpha \frac{\sigma}{\sqrt{N}} \right]$$

avec un niveau de confiance de $(1 - \alpha)\%$, où z_α est le quantile d'une loi normale centrée réduite : $P(|X| \leq z_\alpha) = 1 - \alpha$ où $X \sim N(0, 1)$.

N'oublions pas que la variance σ^2 n'est pas plus connue que l'intégrale cherchée I , ainsi, nous estimons la variance $\sigma^2 = \text{Var}(f(U))$ par :

$$\hat{\sigma}^2 = S_N^2 = \frac{1}{N-1} \sum_{i=1}^N (f(u^i) - I_N)^2.$$

Donc, la particularité très intéressante d'un calcul Monte-Carlo c'est qu'il nous permet à la fois de calculer l'approximation de l'espérance ainsi que sa variance, ce qui permet d'obtenir un intervalle de confiance à 95% comme suit :

$$\left[\frac{1}{N} \sum_{i=1}^N f(u^i) - 1,96 \frac{\hat{\sigma}}{\sqrt{N}}; \frac{1}{N} \sum_{i=1}^N f(u^i) + 1,96 \frac{\hat{\sigma}}{\sqrt{N}} \right].$$

Ça veut dire que l'estimateur de la marge d'erreur de Monte-Carlo $\hat{E} = z_\alpha \frac{\hat{\sigma}}{\sqrt{N}}$ décroît comme $O(\frac{1}{\sqrt{N}})$. Si nous augmentons le nombre de points d'un facteur 100, l'estimateur de la marge d'erreur diminuera de 10 fois.

L'estimation d'erreur est probabiliste. Alors on n'est jamais certain que l'erreur soit inférieure à une certaine limite; il est toujours possible qu'on ne soit pas chanceux avec les échantillons de la méthode de Monte-Carlo. (Thiémard, 2000).

La méthode de Monte-Carlo est largement utilisée parce qu'elle est facile à appliquer et en plus l'erreur ne dépend ni de la dimension n de la variable, ni de la régularité de la fonction f . Cependant les inconvénients de cette méthode est que la vitesse de convergence est moins rapide -seulement de $O(\frac{1}{\sqrt{N}})$ - comparée à

celle de la méthode de quasi-Monte-Carlo, que nous détaillerons dans la prochaine section.

3.2.4 Simulation des variables aléatoires

Une bonne utilisation de la méthode de Monte-Carlo dépend essentiellement d'un bon choix de la qualité de la source de hasard utilisée, car l'utilisation d'un échantillon de points qui nous semble en apparence aléatoire mais qui présente de mauvaises propriétés, peut nous conduire à des résultats erronés. Cette méthode nécessite l'utilisation de l'ordinateur, qui est une machine purement déterministe et non conçue pour engendrer un comportement aléatoire, d'où le recours à la simulation, à savoir, la génération d'un processus ayant uniquement l'apparence aléatoire par des moyens purement déterministes. Nous parlerons alors de générateurs de nombres pseudo-aléatoires.

3.2.4.1 Simulation des variables aléatoires uniformément distribuées sur $[0, 1]$

L'obtention d'une suite de variables aléatoires indépendantes uniformes sur $[0, 1]$ est possible via des langages tel que *Matlab* ou *C++* qui réalisent des appels successifs à un générateur de nombres pseudo-aléatoires. La méthode la plus utilisée est celle des congruences linéaires qui consiste à générer une suite de nombres entiers compris entre 0 et $m - 1$ de la façon suivante : en partant d'une germe quelconque $y_1 \in 0, \dots, m$, la séquence $\{y_1, y_2, \dots\}$ est générée suivant la relation de récurrence :

$$y_i = (ay_{i-1} + c) \bmod m, \quad i = 2, 3, \dots \quad (3.4)$$

avec a , m et c qui représentent des entiers non négatifs appelés respectivement multiplicateur, module et incrément. On parle d'un générateur multiplicatif à congruences linéaires quand $c = 0$. La séquence engendrée par (3.4) est périodique, avec une longueur de période ne dépassant pas m^2 . Le choix de ces entiers

est crucial pour obtenir une suite des points avec de bonnes caractéristiques statistiques. D'une manière générale, on prend souvent :

$$a = 31415821, \quad b = 1 \quad \text{et} \quad m = 10^8.$$

Et on espère que la séquence $\{x_1, x_2, \dots\}$, tel que $x_i = \frac{y_i}{m}$, présente des caractéristiques proches de celles d'une suite de variables aléatoires indépendantes et uniformément distribuées sur $([0, 1])$. (Thiérmard, 2000).

3.2.5 Inversion de la fonction de répartition

L'inversion de la fonction de répartition, proposée initialement par Von Neumann en 1947, est une approche qui permet de simuler une variable X aléatoire à partir de sa fonction de répartition.

Proposition 3.2.1. *Soit U une variable aléatoire de loi uniforme sur $[0, 1]$.*

Soit X une variable aléatoire réelle de fonction de répartition F ;

$F(x) = P(X \leq x)$, dont l'inverse généralisée (appelé quantile) est définie par :

$F^{-1}(u) = \inf\{x : F(x) \geq u\}$. Alors $F^{-1}(U)$ suit la même loi que X .

Inversement, si F est continue, alors $F(X)$ est uniforme sur $[0, 1]$.

Ainsi à partir de cette proposition, nous déduisons que nous pourrions générer des variables des autres lois à partir des uniformes et l'inverse de la fonction de répartition de cette loi. Voici quelques exemples :

- Loi uniforme sur $[a, b]$

Si une variable U suit la loi uniforme sur $[0, 1]$, alors la variable $X = a + (b - a)U$ suit la loi uniforme sur $[a, b]$.

- Loi exponentielle

Si une variable U suit une loi uniforme sur $[0, 1]$, alors $X = -\frac{1}{\lambda} \ln(U)$ suit la loi exponentielle de paramètre λ (noté $Exp(\lambda)$).

Alors pour générer une séquence $(x_i)_{1 \leq i \leq N}$ de N réalisations indépendantes de la variable X qui suit une loi $Exp(\lambda)$, il suffit de générer une séquence $(u_i)_{1 \leq i \leq N}$ de N réalisations indépendantes de la variable U uniformément distribuée sur $[0, 1]$ et nous posons $x_i = -\frac{1}{\lambda} \ln(u_i)$ pour tout i , $1 \leq i \leq N$

- Loi de Bernoulli

Soit $p \in [0, 1]$. Si une variable U suit une loi uniforme sur $[0, 1]$.

Alors si $U < 1 - p$, on pose $X = 0$, sinon, $X = 1$. Alors X suit la loi de Bernoulli de paramètre p .

3.3 Méthode de quasi-Monte-Carlo

3.3.1 Introduction

Dans cette partie de mon mémoire, nous allons présenter une méthode qui est devenue, durant les dernières années, l'intérêt de plusieurs chercheurs, en utilisant comme références la thèse de Pillards (2006) intitulée « Quasi-Monte-Carlo integration over a simplex and the entire space » ainsi que la thèse de Thiémard (2000) qui porte le titre « Sur le calcul et la majoration de la discrédance à l'origine ».

Les méthodes de Monte-Carlo utilisent des points pseudo-aléatoires. Les méthodes de quasi-Monte-Carlo visent à faire mieux. La figure 3.1 (à gauche) montre un ensemble de 100 pseudo-aléatoires sur $[0, 1]^2$. Les points aléatoires forment des grappes, où plus de points se trouvent à proximité les uns aux autres. Il y a aussi certains endroits, où il y a des lacunes, avec relativement de grands espaces vides. La méthode de quasi-Monte-Carlo essaie d'éviter ces lacunes et ces grappes en utilisant plutôt un ensemble de points uniformes. La notion de la discrédance que nous aborderons dans les prochaines sections, mesure l'uniformité de l'ensemble des points. Voir (Pillards, 2006).

Une réaction intuitive est d'utiliser des grilles, comme par exemple la grille 10 par

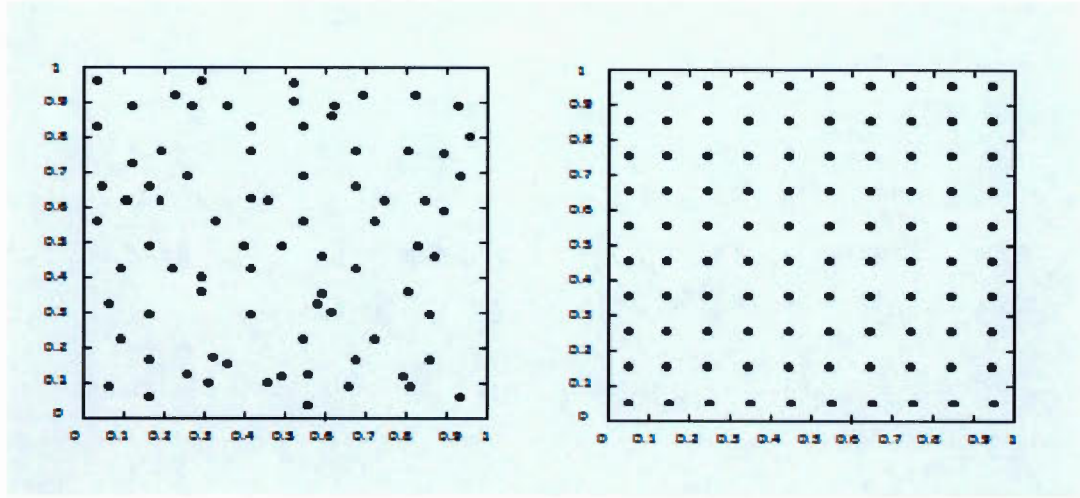


Figure 3.1 À gauche : 100 points aléatoires sur $[0, 1]^2$. À droite : une grille de 10 par 10

10 qui est représentée sur la figure 3.1 (à droite).

Nous avons la possibilité de créer une grille de dimension n en utilisant un ensemble des points uniformes sur $[0, 1]$; $\{x_1, \dots, x_k\}$ avec $0 \leq x_1 < x_2 < \dots < x_k \leq 1$. Alors une grille de dimension n est la combinaison de ces points à toutes les dimensions :

$$S(N) = \{(x_{i_1}, \dots, x_{i_n}) \mid \forall j \in \{1, \dots, n\}, x_{i_j} \in \{x_1, \dots, x_k\}\}, \quad (3.5)$$

où $S(N) \subset [0, 1]^n$ est une grille ou séquence de N points de $[0, 1]^n$ et $N = k^n$.

Une grille semble bien évidemment la meilleure solution. Cependant, une grille présente plusieurs inconvénients. Un premier inconvénient est le nombre de points N qui est toujours de la forme k^n où k est le nombre de points dans l'ensemble des points unidimensionnelles, ce qui signifie qu'il est impossible d'utiliser une grille pour des problèmes de grande dimension (par exemple $n = 100$).

De plus, pour des problèmes de dimension inférieure, le nombre de points N que nous pouvons choisir est très restreint. Par exemple pour $n = 10$, nous pouvons

choisir :

$N \in \{1^{10}, 2^{10}, 3^{10}, 4^{10}, \dots\}$, cela veut dire que :

$N \in \{1, 1024, 59049, 1048576, 9765625, 60466176, \dots\}$.

Un deuxième inconvénient des grilles est que les projections des points sur un ensemble de dimension inférieure tombent au-dessus les uns des autres. Les projections uni-dimensionnelles ne contiennent que k différents points.

La spécificité majeure de la méthode de quasi-Monte-Carlo est le déterminisme, qui se manifeste à deux niveaux :

- premièrement au niveau du choix de l'échantillon déterministe des points utilisés pour approximer l'intégrale sur un cube. L'ensemble de ces points $\{x^1, \dots, x^N\}$ doit constituer une séquence à discrédance faible ;
- deuxièmement au niveau de l'estimation d'erreur d'approximation

$$\left| \int_{[0,1]^n} f(x) dx - \frac{1}{N} \sum_{i=1}^N f(x^i) \right|.$$

Quelques notations constamment utilisées aux paragraphes suivants :

- * $[0, 1)^n$: désigne le cube unité semi-ouvert à droite en dimension n .
- * P : représente un intervalle multidimensionnel semi-ouvert de la forme

$$P = \prod_{j=1}^n [\alpha_j^P, \beta_j^P), \quad \text{ou} \quad 0 \leq \alpha_j^P \leq \beta_j^P \quad \text{pour tout } j \in \{1, \dots, n\}.$$
- * \mathcal{J}_n^* : représente l'ensemble des intervalles inclus dans $[0, 1)^n$ et ancrés à l'origine :

$$\mathcal{J}_n^* = \{P = \prod_{j=1}^n [0, \beta_j^P) : 0 \leq \beta_j^P \leq 1\}.$$

- * $Vol(P)$: est le volume d'un intervalle P : $Vol(P) = \prod_{j=1}^n (\beta_j^P - \alpha_j^P)$.
- * S : représente une séquence de points (c'est-à-dire un ensemble ordonné d'éléments non nécessairement distincts) dénombrable de points que l'on note : $S = \{x_1, x_2, \dots\}$.

Les points considérés appartiennent toujours au cube unité $(0, 1)^n$.

S est appelée une suite dans le cas d'une séquence infinie.

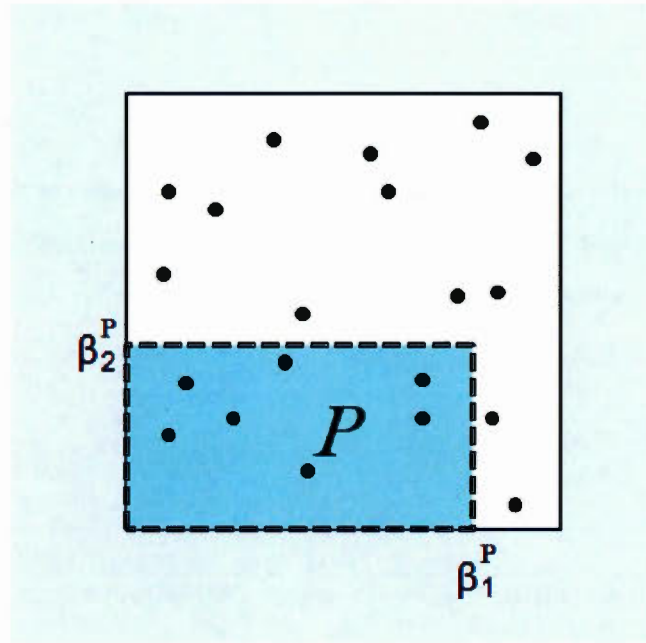


Figure 3.2 Une séquence de $N = 20$ points et un intervalle (rectangle) $P \in \mathcal{J}_2^*$ dans le carré unité $[0, 1]^2$.

- * $S(N)$ (où S est une séquence avec $\text{card}(S) \geq N$) : désigne la sous-séquence constituée des N premiers points de S : $S(N) = \{x_1, \dots, x_N\}$.
- * $A(E, S)$: représente le nombre de points de la séquence S appartenant à l'ensemble $E \subset (0, 1)^n$.
- * $f(N) = O(g(N))$ si et seulement s'il existe une valeur N_0 et une constante réelle c tel que : $\forall N \geq N_0$, on a : $f(N) \geq cg(N)$.

3.3.2 La discrédance

Définition 6. (Pillards, 2006)

Soit une séquence S d'au moins N points dans $[0, 1]^n$. Soit \mathcal{J}_n^* la famille des intervalles $P = \prod_{j=1}^n [0, \beta_j^P)$: $0 \leq \beta_j^P \leq 1$. Soit $A(P, S(N))$ le nombre de points de la séquence $S(N)$ appartenant à l'ensemble P .

La *discr pance* de la s quence S est d finie par :

$$D_N^*(S) = \sup_{P \in \mathcal{J}_n^*} \left| \frac{A(P, S(N))}{N} - \text{Vol}(P) \right|. \quad (3.6)$$

La discr pance $D_N^*(S)$ est d termin e en cherchant l'intervalle ancr    l'origine qui contient la densit  la plus anormalement faible ou le plus anormalement  lev e de points comparativement   son volume. La discr pance est une mesure de la non-uniformit  de S , restreinte   ses N premiers points. Il est clair que

$$0 \leq D_N^*(S) \leq 1$$

et que plus la discr pance est petite, plus la s quence est uniforme.

3.3.3 Calcul de la discr pance d'une s quence de points

Malgr  le fait que le calcul de la discr pance d'une s quence en dimension quelconque soit r put  comme  tant une t che  norm ment difficile, Niederreiter (1992) a pu montrer la possibilit  de minimiser ce probl me et le r soudre en un nombre fini d' tapes. Le seul inconv nient est que la complexit  des diff rents algorithmes connus cro t de mani re exponentielle avec la dimension, ce qui nous oblige souvent de se contenter de bornes inf rieures et sup rieures sur la discr pance.

3.3.3.1 Le cas abordable : la discr pance d'une s quence de dimension 1

Th or me 3.3.1. (*Niederreiter, 1992*)

Pour une s quence de points de dimension 1 ; $S = \{x_1, \dots, x_N\}$,

avec, $0 \leq x_1 \leq x_2 \leq \dots \leq x_N \leq 1$, nous avons :

$$D_N^*(S) = \frac{1}{2N} + \max_{1 \leq i \leq N} \left| x_i - \frac{2i-1}{2N} \right|. \quad (3.7)$$

Alors la discr pance la plus faible sera $D_N^*(S) = \frac{1}{2N}$ si nous prenons $x_i = \frac{2i-1}{2N}$.

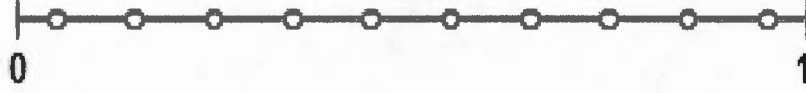


Figure 3.3 Séquence de $N = 10$, points uniformément distribués sur $[0, 1]$.

Par conséquent, la séquence de points optimale de N points définie en termes de la discrédance est :

$$S = \left\{ \frac{1}{2N}, \frac{3}{2N}, \dots, \frac{2N-1}{2N} \right\} \quad (3.8)$$

et sa discrédance est exactement égale à $\frac{1}{2N}$. Voir la figure 3.3.

3.3.3.2 Des séquences des points de faible discrédance

Théorème 3.3.2. (*Schmidt, 1972*)

Il existe une constante $c > 0$ telle que pour toute séquence S dans l'intervalle unité $[0, 1]$, on a :

$$D_N(x) \geq c \frac{\log N}{N} \text{ pour une infinité de valeurs de } N.$$

Définition 7. *Une séquence de points unidimensionnels est une séquence de faible discrédance si :*

$$D_N^*(S) = O\left(\frac{\log(N)}{N}\right) \text{ pour } N \geq 2.$$

Théorème 3.3.3. (*Niederreiter, 1992*)

Il existe une constante $B(n) > 0$, telle que pour toute séquence

$S = \{x_1, \dots, x_N\} \subset [0, 1]^n$, on a :

$$D_N^*(S) \geq B(n) \frac{\log^{n-1}(N)}{N}, \quad (3.9)$$

où la constante $B(n)$ dépend seulement de dimension n .

Définition 8. Nous disons d'une séquence S en dimension $n \geq 1$ est de faible discrétance si :

$$D_N^*(S) = O\left(\frac{\log^n(N)}{N}\right), \quad \text{pour } N \geq 2.$$

3.3.4 Exemples des suites de points à discrétance faible

3.3.4.1 Les séquences de van der Corput

Les suites de van der Corput (1935) sont des suites à discrétance faible dans l'intervalle unité $[0, 1]$.

Soit un entier b tel que $b \geq 2$. Chaque entier naturel $i \in \mathbb{N}$ peut être écrit de manière unique dans la base b comme :

$$i = \sum_{j=0}^{\infty} a_j b^j,$$

avec $a_j \in \mathbb{Z}_b = \{0, 1, \dots, b-1\}$.

On remarque que les coefficients $a_j = 0$ pour tout $j \geq \lfloor \log_b i \rfloor$, alors le nombre de termes de la somme ci-dessus est en fait fini. Nous nous servons de cette représentation pour définir la fonction radicale inverse ϕ_b :

$$\phi_b(i) = \sum_{j=0}^{\infty} a_j b^{-(j+1)}, \quad \text{pour tout } i \in \mathbb{N}. \quad (3.10)$$

On peut montrer facilement que $\phi_b(i) \in [0, 1)$, pour tout $i \in \mathbb{N}$.

Définition 9 (Suite de van der Corput (1935)). Soit un entier $b \geq 2$. La suite de points $S_{van.Cor.} = \{x_0, x_1, \dots\} \subset I$ donnée par $x_i = \phi_b(i)$, est appelée suite de van der Corput en base b .

Exemple : En base $b = 2$, les premiers points de la suite de van der Corput s'obtiennent de manière suivante :

$i = 0$ s'écrit 0 en base 2 donc $x_0 = \phi_b(0) = 0 * 2^{-1} = 0$,
 $i = 1$ s'écrit 1 en base 2 donc $x_1 = \phi_b(1) = 1 * 2^{-1} = 1/2$,
 $i = 2$ s'écrit 10 en base 2 donc $x_2 = \phi_b(2) = 0 * 2^{-1} + 1 * 2^{-2} = 1/4$,
 $i = 3$ s'écrit 11 en base 2 donc $x_3 = \phi_b(3) = 1 * 2^{-1} + 1 * 2^{-2} = 3/4$,
 $i = 4$ s'écrit 100 en base 2 donc $x_4 = \phi_b(4) = 0 * 2^{-1} + 0 * 2^{-2} + 1 * 2^{-3} = 1/8$.

la séquence de van der Corput de base 2 est :

$$S_{van.Cor.}^2 = \{0, 1/2, 1/4, 3/4, 1/8, \dots\}.$$

Le discrédance d'une séquence de Van der Corput $S_{van.Cor.}$, pour $N \geq 2$, est $D_N^*(S) = O(\frac{\log(N)}{N})$, donc cette séquence est de faible discrédance.

3.3.4.2 Les séquences de Halton

Les séquences de Halton (1960) sont des généralisations en dimension quelconque $n \geq 1$ des séquences de van der Corput. L'idée est de considérer la fonction radicale inverse ϕ_b dans différentes bases simultanément. Halton (1960) a démontré que ses séquences sont à discrédance faible.

Définition 10. Soient b_1, \dots, b_n des entiers positifs premiers entre eux deux à deux. La séquence $H_{b_1, \dots, b_n} = \{x_0, x_1, \dots\}$ donnée par :

$$x_i = (\phi_{b_1}(i), \dots, \phi_{b_n}(i)) \in [0, 1]^n$$

est appelée suite de Halton en bases b_1, \dots, b_n .

Ainsi, pour tout $j \in \{1, \dots, n\}$, la suite constituée des composantes $\{x_{j_0}, x_{j_1}, \dots\}$ des points d'une suite de Halton $H_{b_1, \dots, b_n} = \{x_0, x_1, \dots\}$ n'est autre que la suite de van der Corput en base b_j , voir la figure 3.4.

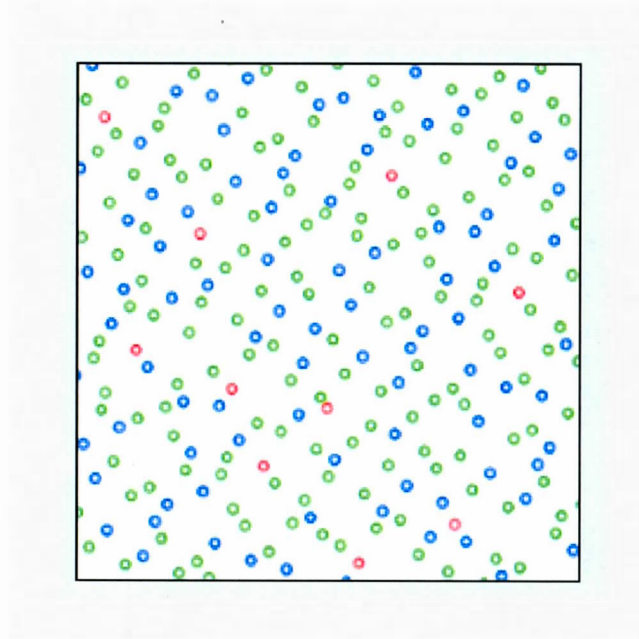


Figure 3.4 Les premiers 256 points de la séquence de Halton en base 2 et 3. Les 10 premiers points de la séquence sont en rouge de 1 à 10, suivis par les points bleus de 11 à 100 ainsi que les points verts de 101 à 256. Figure tirée à partir de https://en.wikipedia.org/wiki/Halton_sequence

3.3.5 Théoreme de Koksma et Hlawka

Le théorème de Koksma et Hlawka est l'un des résultats fondamentaux de la méthode de quasi-Monte-Carlo. Il est souvent désigné par l'inégalité de Koksma-Hlawka. (?) et (?).

Théorème 3.3.4. ? et ?

Si $f : [0, 1]^n \rightarrow \mathbb{R}$ est une fonction à variation $V(f)$ bornée au sens de Hardy et Krause, alors pour toute séquence de points : $S = \{x_1, \dots, x_N\} \subset [0, 1]^n$, avec $N \geq 2$, on a :

$$\left| \frac{1}{N} \sum_{i=1}^N f(x_i) - \int_{[0,1]^n} f(t) dt \right| \leq V(f) D_N^*(S). \quad (3.11)$$

Voir la définition la variation au sens de Hardy et Krause d'une fonction sur la thèse de Pillards (2006), pages 22 à 24.

L'inégalité de Koksma et Hlawka stipule que l'erreur d'approximation de l'intégrale sur $[0, 1]^n$:

$$| \int_{[0,1]^n} f(x)dx - \frac{1}{N} \sum_{i=1}^N f(x^i) |$$

est au maximum égale au produit de la variation $V(f)$ (une grandeur qui reflète l'irrégularité de la fonction f) et de la discrétance $D_N^*(S)$ (qui mesure uniquement la qualité de la répartition des points dans la séquence).

En théorie, la méthode de quasi-Monte-Carlo possède deux atouts majeurs : une convergence asymptotiquement plus rapide que la méthode de Monte-Carlo et la possibilité de calculer des bornes déterministes $V(f) D_N^*(S)$ sur l'erreur (3.11). Malgré le fait que, l'exploitation de la seconde propriété soit encore un vœu pieux, car la discrétance et la variation au sens de Hardy et Krause sont très difficiles à calculer, mais, sa rapidité de convergence séduit un bon nombre d'anciens utilisateurs de la méthode de Monte-Carlo.

CHAPITRE IV

COMPARAISON ENTRE LES TROIS MÉTHODES DANS LE CALCUL DE LA VRAISEMBLANCE BASÉE SUR LES RANGS

4.1 Introduction

La valeur exacte de la vraisemblance d'une copule basée sur les rangs est difficile à obtenir de façon théorique. Pour approximer cette valeur, nous aurons donc recours à l'ordinateur en utilisant un programme puissant et rapide puisqu'il s'agit de calculs intensifs. Le premier objectif est de savoir jusqu'à quel rang, ou quelle dimension, l'ordinateur sera capable d'exécuter ce calcul via ce programme. Le deuxième objectif est de comparer la performance des trois méthodes utilisées. Pour ce faire, nous utiliserons *C++* qui est un langage de programmation compilé rapide et performant par rapport aux autres langages connus.

4.2 Calcul de l'approximation de la vraisemblance basée sur les rangs avec les trois méthodes

4.2.1 Méthode de Monte-Carlo

Nous utiliserons la formule de la proposition 2.5.1 vu au chapitre II :

Soient (X_i, Y_i) , $i = 1, \dots, n$, des bivariées indépendantes et identiquement distribuées ayant une fonction de répartition $H \in \mathcal{H}$ avec une copule C , et leurs

fonctions de répartition marginales (continues), sont respectivement F et G .

Si la copule C a une fonction de densité c , nous avons donc :

$$P(S(X, Y) = s) = \frac{1}{n!} E \left(\prod_{i=1}^n c(U_{(i)}, V_{(s(i))}) \right), \quad (4.1)$$

où $U_{(1)}, \dots, U_{(n)}$ et $V_{(1)}, \dots, V_{(n)}$ sont respectivement des statistiques d'ordre de U_1, \dots, U_n et V_1, \dots, V_n , deux échantillons de variables aléatoires indépendants de la distribution uniforme sur $(0, 1)$.

Pour appliquer la méthode de Monte-Carlo dans le calcul de la vraisemblance en utilisant la formule (4.1), nous suivrons les étapes suivantes :

– **Simulation des statistiques d'ordre** $U_{(1)}, \dots, U_{(n)}$ et $V_{(1)}, \dots, V_{(n)}$

Puisque la commande `-sort-` ne se trouve pas en $C++$ et pour alléger le plus possible le programme, nous procèderons ainsi :

Nous générerons un échantillon d'observations x_1, \dots, x_{n+1} de variables aléatoires X_1, \dots, X_{n+1} indépendantes et uniformément distribuées sur $(0, 1)$.

Ensuite nous générerons d'observations y_1, \dots, y_{n+1} tels que

$y_i = -\log(x_i)$ pour $i = 1, \dots, n+1$. Cela signifie que y_1, \dots, y_{n+1} sont des simulés de variables aléatoires Y_1, \dots, Y_{n+1} indépendantes et identiquement distribuées de loi $\exp(\lambda = 1)$.

$$\text{Pour } i = 1, \dots, n, \text{ nous posons } u_{(i)} = \frac{\sum_{k=1}^i y_k}{\sum_{l=1}^{n+1} y_l}.$$

Alors $u_{(1)}, \dots, u_{(n)}$ sont des observations de statistiques d'ordre $U_{(1)}, \dots, U_{(n)}$ d'un échantillon de variables aléatoires U_1, \dots, U_n indépendantes et distribuées uniformément sur $(0, 1)$.

Nous générerons de la même façon des observations $v_{(1)}, \dots, v_{(n)}$ de statistiques

d'ordre $V_{(1)}, \dots, V_{(n)}$ d'un échantillon de variables aléatoires V_1, \dots, V_n indépendantes et distribuées uniformément sur $(0, 1)$.

Soit s une permutation quelconque de S_n et M_s sa matrice. Alors,

$$(v_{(s(1))}, \dots, v_{(s(n))})^T = M_s(v_{(1)}, \dots, v_{(n)})^T.$$

– Calcul de l'approximation de la vraisemblance

Nous utiliserons les deux échantillons indépendants des statistiques d'ordres générés ci-dessus pour approximer la valeur de la vraisemblance :

$$\begin{aligned} n! P(S(X, Y) = s) &= E \left(\prod_{i=1}^n c(U_{(i)}, V_{(s(i))}) \right), \\ &\approx vr.MC_s(n, \theta, N), \end{aligned}$$

tel que $vr.MC_s(n, \theta, N)$ est l'approximation de la vraisemblance par la méthode de Monte-Carlo multipliée par $n!$. Elle est en fonction de la permutation s de taille n , de paramètre de la copule θ et de la taille de l'échantillon des points simulés N .

Nous programmerons la fonction $vr.MC_s(n, \theta, N)$ sur $C++$ en utilisant la formule : $vr.MC_s(n, \theta, N) \approx E \left(\prod_{i=1}^n c(U_{(i)}, V_{(s(i))}) \right)$.

Selon la loi des grands nombres, il faudrait prendre une grande taille d'échantillon N , pour que cette approximation soit relativement proche de la valeur exacte.

4.2.2 Méthode de quasi-Monte-Carlo

Dans cette méthode, nous utiliserons la même formule de la proposition 2.5.1 et nous procéderons de la même façon que pour la méthode de Monte-Carlo, mais le choix de l'échantillon des points sera différent.

Il s'agit alors de remplacer l'approximation de Monte-Carlo dans laquelle les points sont simulés de variables aléatoires indépendantes et identiquement distribuées sur $(0, 1)$, par une approximation de quasi-Monte-Carlo dans laquelle les points de l'échantillon sont choisis d'une façon déterministe et constituent une séquence à faible discrédance.

Les échantillons choisis dans cette méthode de quasi-Monte-Carlo seront les séquences de Halton qui sont des généralisations des séquences de van der Corput, vues au chapitre précédent, dans la section 3.3.4.

Nous choisirons alors b_1, \dots, b_n , des entiers positifs premiers entre eux, deux à deux. La séquence $H_{b_1, \dots, b_n} = \{x_0, x_1, \dots\}$ donnée par :

$$x_i = (\phi_{b_1}(i), \dots, \phi_{b_n}(i)) \in [0, 1]^n$$

est la suite de Halton en bases b_1, \dots, b_n .

Ainsi, pour tout $j \in \{1, \dots, n\}$, la suite constituée des j^e composantes $\{x_{j_0}, x_{j_1}, \dots\}$ des points d'une suite de Halton n'est que la suite de van der Corput en base b_j , et ϕ_{b_i} est la fonction radicale inverse de la base b_i , pour $i = 1, \dots, n$.

Nous prendrons les N premiers points de cette séquence :

$$H_{b_1, \dots, b_n}(N) = \{x_0, x_1, \dots, x_{N-1}\}.$$

Nous utilisons cette séquence de points pour calculer une approximation de la vraisemblance via la formule ci-dessous :

$$\begin{aligned} n! P(S(X, Y) = s) &= E \left(\prod_{i=1}^n c(U_{(i)}, V_{(s(i))}) \right), \\ &\approx vr.QMC_s(n, \theta, N), \end{aligned}$$

tel que $vr.QMC_s(n, \theta, N)$ est l'approximation de la vraisemblance par la méthode de quasi-Monte-Carlo multipliée par $n!$.

Nous programmerons la fonction $vr.QMC_s(n, \theta, N)$ sur $C++$ qui dépend de la permutation $s \in S_n$, le paramètre de la copule θ et la taille de la séquence N .

4.2.3 Méthode de Grundmann-Möller

Dans cette méthode, nous exploiterons le théorème de Grundmann-Möller (1978) de l'intégration sur le simplexe de dimension n par le théorème 3.1.3 :

Soient $r \in \mathbb{N}$, $n \in \mathbb{N}$ et $d = 2r + 1$. Donc

$$\sum_{i=0}^r (-1)^i 2^{-2r} \frac{(d+n-2i)^d}{i!(d+n-i)!} \sum_{|\beta|=r-i}^{\beta_0 \geq \dots \geq \beta_n} f\left(\frac{2\beta_0+1}{d+n-2i}, \dots, \frac{2\beta_n+1}{d+n-2i}\right) = I_n f - R_{nd} f \quad (4.2)$$

est une formule d'intégration invariante de degré d , ce qui signifie que l'erreur $R_{nd} f = 0$ pour toute fonction polynomiale f telle que $f \in \mathcal{P}_d$, où \mathcal{P}_d est l'espace linéaire des monômes de degrés au plus d .

Ainsi, nous exploiterons la lemme 3.1.4 :

Pour tout $\eta \in \mathbb{N}^n$ tel que $|\eta| = |\eta_1 + \dots + \eta_n| \leq d$, nous avons $R_{nd} x^\eta = 0$.

Car le $x^\eta = (x_1, \dots, x_n)^\eta \in \mathcal{P}_d$.

Le théorème 3.1.3 et la lemme 3.1.4 nous permettront de trouver la valeur exacte de la vraisemblance d'une copule polynomiale basée sur les rangs, ou de l'approximer dans le cas des copules non polynomiales, via la formule déjà vue à la proposition 2.5.2 :

$$P(S(X, Y) = s) = n! \int_{\Delta_n^2} c(\varphi_s(x, y)) dx dy, \quad (4.3)$$

où φ_s est une bijection définie par, $\varphi_s : \Delta_n \times \Delta_n \rightarrow E_s$

$$\varphi_s(x, y) = (Lx, M_s Ly),$$

avec $E_s = \{(u, v) \in [0, 1]^{2n} : R(u, v) = (e, s)^\top\}$, L est une matrice triangulaire inférieure et M_s est une matrice de permutation correspondant à $s \in S_n$.

Pour calculer la vraisemblance par un algorithme en $C++$ via la méthode de Grundmann-Möller, nous procéderons alors ainsi :

- Nous déterminerons tous les points (nœuds) appartenant à des classes d'équivalences sous forme : $((\frac{2\beta_0+1}{d+n-2i}, \dots, \frac{2\beta_n+1}{d+n-2i}))$. Ainsi, nous trouverons le poids associé à chaque nœud sous forme : $(-1)^i 2^{-2r} \frac{(d+n-2i)^d}{i!(d+n-i)!}$ pour tout $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$ tel que $|\beta| = r - i$, $\beta_0 \geq \dots \geq \beta_n$ et $i = 0, \dots, r$, avec r un entier naturel qui est fixé à l'avance et $d = 2r + 1$.

Plus nous augmenterons t , plus nous disposerons de points (nœuds) à utiliser pour calculer l'intégrale de Grundmann-Möller, ainsi l'erreur $|R_{nd} f|$ de la formule de l'intégration invariante diminue.

- Nous utiliserons le premier coté de la formule (4.2) :

$$\sum_{i=0}^r (-1)^i 2^{-2r} \frac{(d+n-2i)^d}{i!(d+n-i)!} \sum_{|\beta|=r-i}^{\beta_0 \geq \dots \geq \beta_n} f\left(\left(\frac{2\beta_0+1}{d+n-2i}, \dots, \frac{2\beta_n+1}{d+n-2i}\right)\right)$$

deux fois pour avoir une approximation de l'intégrale double sur le simplexe Δ_n ; $\int_{\Delta_n^2} c_\theta(\varphi_s(x, y)) dx dy$.

Nous calculerons donc la vraisemblance via un algorithme en $C++$ décrivant la fonction $vr.GM_s(n, \theta, d)$ tel que,

$$\begin{aligned} n! P(S(X, Y) = s) &= (n!)^2 \int_{\Delta_n^2} c_\theta(\varphi_s(x, y)) dx dy, \\ &\approx vr.GM_s(n, \theta, d). \end{aligned}$$

tel que, $vr.GM_s(n, \theta, d)$ est l'approximation de la vraisemblance par la méthode de Grundmann-Möller. Elle est en fonction du paramètre de la copule θ , de la taille n de la permutation s et du degré d'intégration d où $d = 2r + 1$.

Remarque 4.2.1. *L'avantage de la méthode de Grundmann-Möller est qu'elle calcule la valeur exacte de la vraisemblance de toute copule polynomiale dont la densité est une combinaison linéaire des monômes x^η tel que $|\eta| \leq d$, cela d'après la lemme 3.1.4.*

Nous pouvons alors démontrer facilement que pour toute copule polynomiale C_θ de degré deux, nous avons :

$$n! P(S(X, Y) = s) = vr.GM_s(n, \theta, d), \text{ pour tout } d = 2r + 1 \geq 3, \quad (4.4)$$

tel que d est le degré de la formule de l'intégration de Grundmann-Möller.

Et, pour toute copule polynomiale C_θ de degré trois, nous avons :

$$n! P(S(X, Y) = s) = vr.GM(n, \theta, d), \text{ pour tout } d = 2r + 1 \geq 5. \quad (4.5)$$

Remarque 4.2.2. *Puisque l'on connaît la valeur exacte de la vraisemblance d'une copule polynomiale à partir de la méthode de Grundmann-Möller, nous sommes à même de constater quelle méthode, entre Monte-Carlo et quasi-Monte-Carlo, est la plus efficace.*

4.3 Comparaison entre les trois méthodes

4.3.1 La valeur exacte de la vraisemblance via la méthode de Grundmann-Möller

4.3.1.1 Copules polynomiales de degré deux

Pour un paramètre $\theta \in [-1, 1]$, $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ est une copule polynomiale de degré deux, de Farlie-Gumbel-Morgenstern.

Elle appartient à une famille paramétrique, et sa densité associée est :

$$c_\theta(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v) = 1 + \theta(1 - 2u)(1 - 2v).$$

Via un algorithme fait sur $C++$, nous pourrions calculer la vraisemblance de la copule polynomiale de degré deux ci-dessus basée sur les rangs.

Tableau 4.1 Calcul de la vraisemblance $n!P(S(X, Y) = s)$ de la copule polynomiale de degré deux : $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ par la méthode de Grundmann-Möller pour $n = 4$ et 6 , et $\theta = 0, 8$.

d	n=4	n=6
	vr.GM(n,0.8, d)=n!P(S(X,Y)=s)	
1	0,8278973	0,9927833
3	0,802757	0,9645162
5	0,8027349	0,9641685
7	0,8027349	0,9641684
9	0,8027349	0,9641684
11	0,8027349	0,9641684
13	0,8027349	0,9641684
15	0,8027349	0,9641684
17	0,8027349	0,9641684

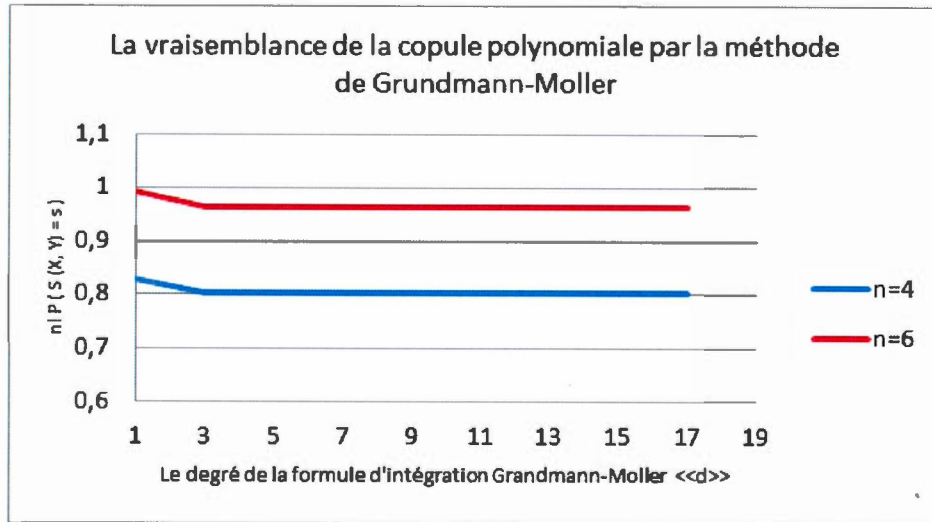


Figure 4.1 La valeur de la vraisemblance $n!P(S(X, Y) = s)$ de la copule polynomiale de degré deux : $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ selon le degré d .

Alors nous calculons :

$$vr.GM(n, \theta, d) \approx (n!)^2 \int_{\Delta_n^2} c(\varphi_s(x, y)) dx dy = n!P(S(X, Y) = s).$$

Dans notre première expérience, nous essayons de vérifier si les résultats de nos calculs appuient la formule (4.4), et nous tentons de trouver la valeur exacte de la vraisemblance pour une permutation s et une copule de paramètre θ .

Dans le tableau 4.1, nous avons calculé $vr.GM_s(n, \theta, d)$ la vraisemblance d'une copule polynomiale d'ordre deux, selon d le degré d'intégration invariante de Grundmann-Möller, pour $n = 4$ et $n = 6$ et $\theta = 0.8$.

La figure 4.1 montre la variation de la fonction $vr.GM$ selon d . Nous remarquons que $vr.GM(n, \theta, d)$ est constante pour tout $d \geq 3$ et cela confirme le théorème d'intégration invariante de Grundman-Möller et la lemme 3.1.4. Alors 0.7108551 et 0.6361141 sont respectivement les valeurs exactes des vraisemblances cherchées multipliées par $n!$, pour $n = 4$ et $n = 6$.

Cette valeur exacte nous permettra donc de comparer l'approximation de la vraisemblance basée sur les rangs via les deux autres méthodes, Monte-Carlo et quasi-Monte-Carlo.

4.3.1.2 Copules polynomiales de degré trois

La famille de copules asymétriques avec des sections cubiques (qui sont des copules polynomiales de degré trois) sont définies par :

$$C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b],$$

tel que : $(a, b) \in \mathbb{R}^2$, $a \neq b$, $|b| \leq 1$ et $[b - 3 - (9 + 6b - 3b^2)^{1/2}]/2 \leq a \leq 1$.

Et la copule densité associée est :

$$\begin{aligned} c_{a,b}(u, v) &= \frac{\partial^2}{\partial u \partial v} C_{a,b}(u, v), \\ &= 1 + (a-b)(1-4u+3u^2)(2v-3v^2) + b(1-2u)(1-2v). \end{aligned}$$

Tableau 4.2 Calcul de la vraisemblance $n!P(S(X,Y) = s)$ de la copule $C_{a,b}(u,v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$ avec la méthode de Grundmann-Möller, avec $a = -1.2$ et $b = 0.6$.

d	n=4	n=6
	vr.GM(-1.2, 0.6, n, d)=n!P(S(X,Y)=s)	
1	0,7303718	0,6269161
3	0,7109904	0,632135
5	0,7109587	0,6361906
7	0,7108552	0,6361184
9	0,7108551	0,6361142
11	0,7108551	0,6361141
13	0,7108551	0,6361141
15	0,7108551	0,6361141

Nous calculons :

$$vr.GM(n, a, b, d) \approx n!P(S(X,Y) = s) = (n!)^2 \int_{\Delta_n^2} c(\varphi_s(x,y)) dx dy.$$

Dans le tableau 4.2 ci-dessus, nous avons calculé $vr.GM(n, a, b, d)$ selon d le degré d'intégration invariante de Grundmann-Möller pour $n = 4$ et $n = 6$, avec $a = -1.2$ et $b = 0.6$.

En ce qui concerne le calcul de la vraisemblance d'une copule polynomiales de degré trois, nous obtenons les mêmes remarques que celles de degré deux, via la méthode de Grundmann-Möller.

En effet, d'après la figure 4.2, nous remarquons que $vr.GM(n, a, b, d)$ est constante pour tout $d \geq 5$. Les résultats obtenus confirment à nouveau la lemme (3.1.4) et le théorème de la formule d'intégration invariante de Grundmann-Möller.

Alors 0.7108551 et 0.6361141 sont respectivement les valeurs exactes des vraisemblances cherchées multipliées par $n!$ pour $n = 4$ et $n = 6$.

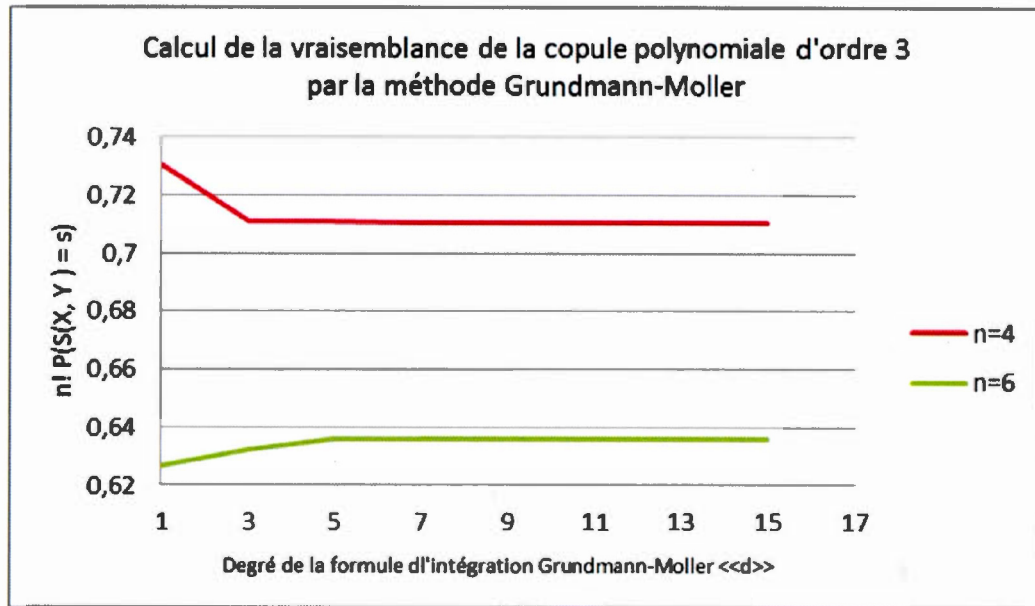


Figure 4.2 L'approximation de la vraisemblance $n! P(S(X, Y) = s)$ selon d avec $C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$ et $a = -1.2$ et $b = 0.6$.

4.3.2 L'approximation par la méthode de Monte-Carlo

4.3.2.1 Copules polynomiales de degré deux

Dans cette section, nous examinons la variation de l'approximation de la vraisemblance par la méthode de Monte-Carlo selon N la taille de l'échantillon des points utilisé dans cette approximation.

Nous calculons via $C++$ la fonction $vr.MC_s$ qui représente l'approximation de la vraisemblance par la méthode de Monte-Carlo. Nous avons :

$$vr.MC_s(n, \theta, N) \approx n! P(S(X, Y) = s) = E \left(\prod_{i=1}^n c(U_{(i)}, V_{(s(i))}) \right),$$

pour différentes tailles de l'échantillon N , pour $n = 4$ et $\theta = 0.8$. Et nous la comparons avec la valeur exacte trouvée par la méthode de Grundmann-Möller ;

vr.GM. Voir le tableau 4.3 ci-dessous :

Tableau 4.3 L'approximation de la vraisemblance $n! P(S(X, Y) = s)$ tel que $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$ par la méthode de Monte-Carlo, selon la taille de l'échantillon N . Pour un $\theta = 0.8$ et $n = 4$.

N	vr.GM	vr.MC
300	0,8027349	0,8223769
600	0,8027349	0,8025798
900	0,8027349	0,8117497
1200	0,8027349	0,8185609
1500	0,8027349	0,8166403
1800	0,8027349	0,817122
2100	0,8027349	0,8168096
2400	0,8027349	0,8152814
2700	0,8027349	0,8135097
3000	0,8027349	0,8150129
3300	0,8027349	0,8138494
3600	0,8027349	0,8110942
3900	0,8027349	0,8103981
4000	0,8027349	0,8110348
6000	0,8027349	0,8068813
8000	0,8027349	0,8066292
10000	0,8027349	0,806553
12000	0,8027349	0,8076638
14000	0,8027349	0,8053858
16000	0,8027349	0,8042062
18000	0,8027349	0,8040502

4.3.2.2 Copules polynomiales de degré trois

Nous procéderons de la même manière pour la copule polynomiale de degré trois :

$C_{a,b}(u, v) = uv + uv(1 - u)(1 - v)[(a - b)v(1 - u) + b]$ pour qu'on puisse tirer, d'une façon générale, une conclusion sur la qualité de l'approximation de la vraisemblance par la méthode de Monte-Carlo selon la taille de l'échantillon N .

Nous calculons via $C++$ la fonction :

$$vr.MC_s(n, a, b, N) \approx n! P(S(X, Y) = s) = E \left(\prod_{i=1}^n c(U_{(i)}, V_{(s(i))}) \right)$$

pour différentes tailles de l'échantillon N . Voir le tableau 4.4.

Tableau 4.4 L'approximation de la vraisemblance $n! P(S(X, Y) = s)$ selon la taille de l'échantillon N , par la méthode de Monte-Carlo, tel que $C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$, où $a = -1.2$, $b = 0.6$ et $n = 4$.

N	vr.GM	vr.MC
300	0,7108551	0,6984249
600	0,7108551	0,6902286
900	0,7108551	0,6943253
1200	0,7108551	0,6949931
1500	0,7108551	0,6954301
1800	0,7108551	0,7004374
2100	0,7108551	0,7019833
2400	0,7108551	0,7051812
2700	0,7108551	0,7073327
3000	0,7108551	0,7076014
4000	0,7108551	0,7061766
6000	0,7108551	0,7061307
8000	0,7108551	0,7076684
10000	0,7108551	0,7089394
12000	0,7108551	0,7101378
14000	0,7108551	0,7109191
16000	0,7108551	0,7090662
18000	0,7108551	0,7088337
20000	0,7108551	0,7090317

Les courbes des figures 4.3 et 4.4 représentant les résultats obtenus dans les tableaux 4.3 et 4.4 montrent que :

- En augmentant la taille de l'échantillon N , l'approximation de la vraisemblance par la méthode de Monte-Carlo « $vr.MC$ » approche relativement de la valeur exacte : $n! P(S(X, Y) = s) = vr.GM(4, 0.8, 3) = 0.8027349$

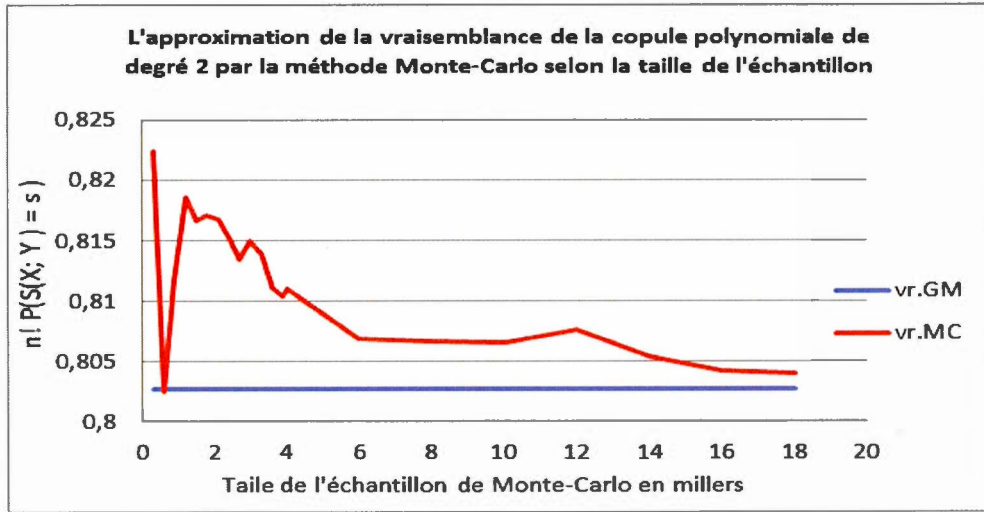


Figure 4.3 La variation de l'approximation $n!P(S(X,Y) = s)$ selon la taille de l'échantillon N , par rapport à la valeur exacte $vr.GM$. Tel que $C_\theta(u,v) = uv(1 + \theta(1-u)(1-v))$, $\theta = 0.8$ et $n = 4$.

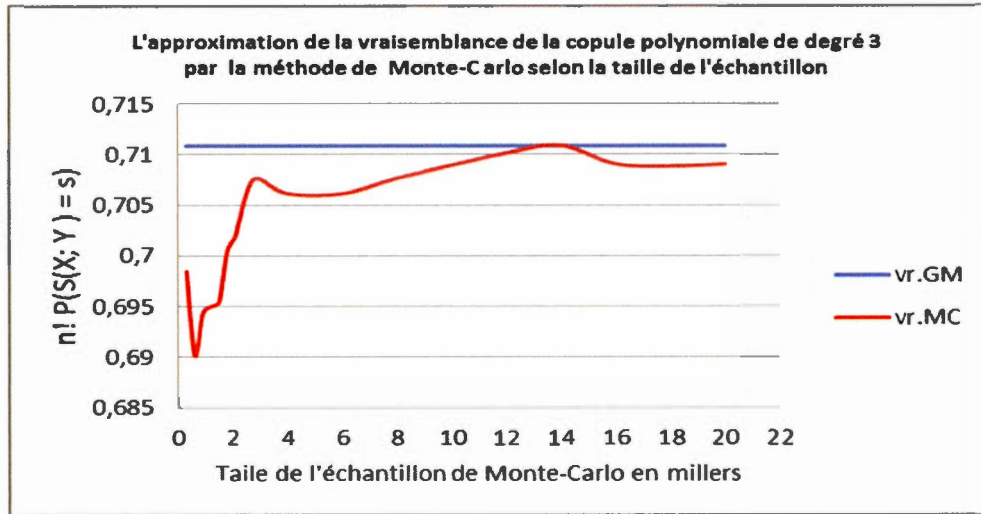


Figure 4.4 La variation de l'approximation $n!P(S(X,Y) = s)$ selon la taille de l'échantillon N , par rapport à la valeur exacte $vr.GM$ $v + uv(1-u)(1-v)[(a-b)v(1-u) + b]$, $n = 4$, $a = -1.2$ et $b = 0.6$.

obtenue par la méthode de Grundmann-Möller.

- Quand N est petit (moins de 8000 points environ), l'approximation de la vraisemblance fluctue davantage et elle n'est pas relativement proche de la valeur exacte 0.8027349.
- Quand N est grand (plus de 80000 points environ), l'approximation de la vraisemblance fluctue moins et elle est relativement proche de la valeur exacte.

4.3.3 Comparaison entre les deux méthodes Monte-Carlo et quasi-Monte-Carlo selon la taille de l'échantillon

4.3.3.1 Copules polynomiales de degré deux

Dans cette section, nous vérifierons laquelle des deux méthodes, Monte-Carlo ou quasi-Monte-Carlo, réussit le mieux dans l'approximation de la vraisemblance pour différentes tailles de l'échantillon N .

Pour une permutation donnée $s \in S_n$ et un paramètre θ de la copule, nous calculons l'approximation de la vraisemblance $vr.MC(n, \theta, N)$ et $vr.QMC(n, \theta, N)$ pour différentes tailles de l'échantillon N et nous les comparons avec la valeur exacte $vr.GM$ obtenue par la méthode de Grundmann-Möller.

Tableau 4.5 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ selon N la taille des échantillons utilisés dans les deux méthodes Monte-Carlo et quasi-Monte-Carlo. On a $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$, $\theta = 0.8$ et $n = 4$.

N	vr.GM	vr.MC	vr.QMC
300	0,8027349	0,8223769	0,8044506
600	0,8027349	0,8025798	0,803932
900	0,8027349	0,8117497	0,8050598
1200	0,8027349	0,8185609	0,806706
1500	0,8027349	0,8166403	0,8057344
1800	0,8027349	0,817122	0,8058578
2100	0,8027349	0,8168096	0,8051298
2400	0,8027349	0,8152814	0,8046533
2700	0,8027349	0,8135097	0,8050609
3000	0,8027349	0,8150129	0,8049007
3300	0,8027349	0,8138494	0,8043947
3600	0,8027349	0,8110942	0,8046058
3900	0,8027349	0,8103981	0,8043331

4.3.3.2 Copules polynomiales de degré trois

Nous procéderons de la même manière pour la copule polynomiale de degré trois afin de pouvoir tirer une conclusion un peu plus générale. Pour une copule $C_{a,b}(u, v) = uv + u v (1 - u) (1 - v)[(a - b) v (1 - u) + b]$, une permutation s et des paramètres a et b , nous calculons $vr.MC(n, a, b, N)$ et $vr.QMC(n, \theta, N)$ pour différentes tailles N de l'échantillon.

Tableau 4.6 L'approximation de la vraisemblance $n! P(S(X, Y) = s)$ de la copule $C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$ avec les trois méthodes. Pour $a = -1.2$ et $b = 0.6$.

N	vr.GM	vr.MC	vr.QMC
300	0,7108551	0,6984249	0,7087696
600	0,7108551	0,6902286	0,7102154
900	0,7108551	0,6943253	0,7095929
1200	0,7108551	0,6949931	0,7118049
1500	0,7108551	0,6954301	0,7118279
1800	0,7108551	0,7004374	0,7121885
2100	0,7108551	0,7019833	0,7111899
2400	0,7108551	0,7051812	0,7107729
2700	0,7108551	0,7073327	0,7114623
3000	0,7108551	0,7076014	0,7118346

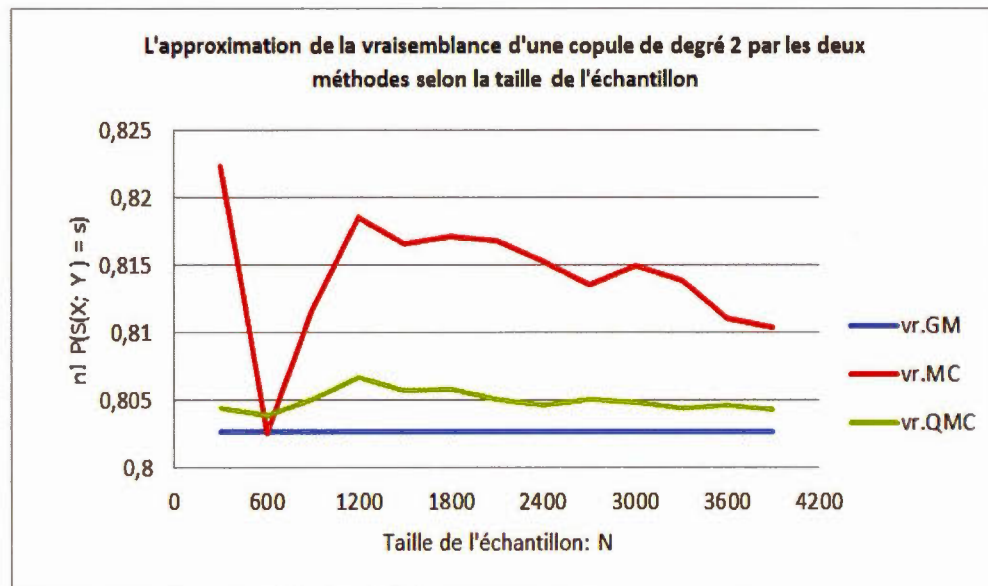


Figure 4.5 La variation de l'approximation de la vraisemblance $n! P(S(X, Y) = s)$ par les deux méthodes Monte-Carlo et quasi-Monte-Carlo selon la taille de l'échantillon N . Avec $C_{\theta}(u, v) = uv(1 + \theta(1-u)(1-v))$ avec $\theta = 0.8$ et $n = 4$.

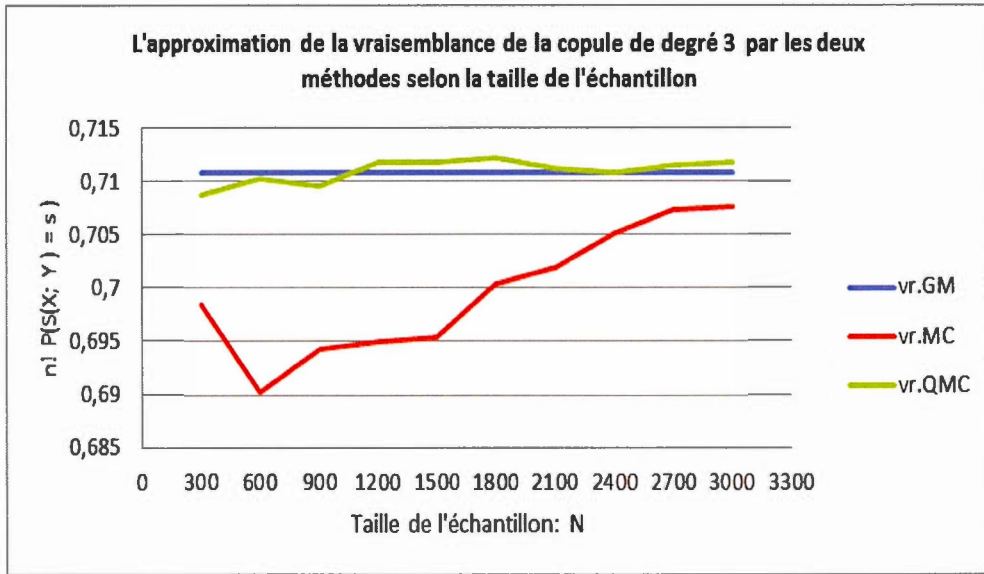


Figure 4.6 La variation de l'approximation de la vraisemblance $n!P(S(X, Y) = s)$ de la copule $C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$ selon la taille de l'échantillon N par les deux méthodes. Avec $a = -1.2$, $b = 0.6$ et $n = 4$.

D'après les deux courbes des figures 4.5 et 4.6, nous remarquons que :

- Pour une permutation donnée s de la taille $n = 4$, l'approximation de la vraisemblance par la méthode de quasi-Monte-Carlo «*vr.QMC*» est toujours proche de la valeur exacte : $n!P(S(X, Y) = s) = vr.GM(4, 0.8, 3) = 0.8027349$ pour toute taille d'échantillon (séquence de Halton), et ce, même pour des petites tailles d'environ 300 ou 400 points.
- Pour la méthode de Monte-Carlo, une grande taille de l'échantillon de points est requise (environ 2700 points ou plus) pour que l'approximation soit proche de la valeur exacte de la vraisemblance.
- L'approximation par la méthode de Monte-Carlo varie davantage et elle n'est pas proche de la valeur exacte, surtout si la taille de l'échantillon est petite, contrairement à la méthode de quasi-Monte-Carlo, pour laquelle

l'approximation varie peu et est plutôt proche de la valeur exacte.

Cela démontre que l'échantillon déterministe (séquence de Halton) utilisé dans la méthode de quasi-Monte-Carlo est plus avantageux pour le calcul de l'approximation de la vraisemblance basée sur les rangs.

Les mêmes remarques ont été obtenues en ce qui a trait à la comparaison entre les deux méthodes selon la taille de l'échantillon, quand nous avons pris une autre permutation s , ainsi d'autres dimensions relativement petites, tel $n = 5, 6, \dots, 10$.

4.3.4 Comparaison entre les trois méthodes selon la taille de la permutation

4.3.4.1 Copules polynomiales de degré deux

Dans cette section, nous vérifierons laquelle des deux méthodes, Monte-Carlo ou quasi-Monte-Carlo, calcule le mieux l'approximation de la vraisemblance quand nous varions n la taille de la permutation s et nous fixons les autres paramètres, comme $\theta = 0.8$, la taille de l'échantillon $N = 3000$ et $C_\theta(u, v) = uv(1 + \theta(1 - u)(1 - v))$.

Nous calculons alors $vr.MC(n, \theta, N)$ et $vr.QMC(n, \theta, N)$, ainsi que la valeur exacte de la vraisemblance $vr.GM(n, \theta, d)$ pour différentes dimensions n . Voir le tableau 4.7.

Tableau 4.7 L'approximation de la vraisemblance $n!P(S(X,Y) = s)$ selon la taille de la permutation n , par les trois méthodes, où $C_\theta(u,v) = uv(1 + \theta(1 - u)(1 - v))$, $\theta = 0.8$, et la taille de l'échantillon $N_{MC} = N_{QMC} = 3000$.

n	vr.GM	vr.MC	vr.QMC
2	0,8222222	0,8218256	0,821966
4	0,8027349	0,8150129	0,8049007
6	0,9641684	0,9582816	0,9660278
8	1,2699973	1,2759383	1,2741269
10	1,7595415	1,7580318	1,76683
12	2,5150578	2,4709358	2,5176366
14	3,6706396	3,6653899	3,6705273
16	5,4364243	5,4212362	5,4271927
18	8,1389749	8,1514723	8,108202
20	12,2850933	12,2190932	12,1561795
22	18,6616706	18,6997633	18,4064287

Nous calculons ensuite l'erreur en pourcentage de l'approximation de la vraisemblance pour les deux méthodes, Monte-Carlo et quasi-Monte-Carlo,

$$Erreur.M.C(n, \theta, N) = \frac{|vr.MC - vr.GM|}{vr.GM} \times 100,$$

$$Erreur.q.M.C(n, \theta, N) = \frac{|vr.QMC - vr.GM|}{vr.GM} \times 100,$$

et nous traçons la courbe de l'erreur de l'approximation pour les deux méthodes en fonction de la dimension n . Voir la figure 4.7 ci-dessous.

4.3.4.2 Copule polynomiale de degré trois

Nous appliquons le même procédé pour la copule polynomiale de degré trois $C_{a,b}(u,v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$ avec $a = -1.2$, $b = 0.6$ et $N_{MC} = N_{qMC} = 3000$, afin de pouvoir tirer des remarques communes entre le calcul de l'approximation de la vraisemblance des deux copules. Voir le tableau 4.8 ci-dessous.

Tableau 4.8 L'approximation de la vraisemblance $n!P(S(X, Y) = s)$ selon n la taille de la permutation s , par les trois méthodes. La copule $C_{a,b}(u, v) = uv + uv(1-u)(1-v)[(a-b)v(1-u) + b]$, $a = -1.2$, $b = 0.6$ et $N_{MC} = N_{qMC} = 3000$.

n	vr.GM	vr.MC	vr.QMC
2	0,9642667	0,9733999	0,9641985
4	0,7109587	0,7076014	0,7118346
6	0,6361906	0,6390708	0,6368584
8	0,6694832	0,6890708	0,6703017
10	0,7703791	0,7902417	0,769436
12	0,9326529	0,9187977	0,9301467
14	1,1635885	1,1688268	1,1560237
16	1,4803122	1,522048	1,4625347
18	1,9083446	1,8780062	1,8747521
20	2,4833119	2,4604399	2,4091091
22	3,2537748	3,2975375	3,1316663
24	4,2853982	4,1243841	4,0767535

Nous calculons ensuite l'erreur en pourcentage de l'approximation de la vraisemblance pour les deux méthodes, Monte-Carlo et quasi-Monte-Carlo, par rapport à la valeur exacte, et nous traçons la courbe de ces deux erreurs d'approximation en fonction de la dimension n . Voir la figure 4.8 ci-dessous.

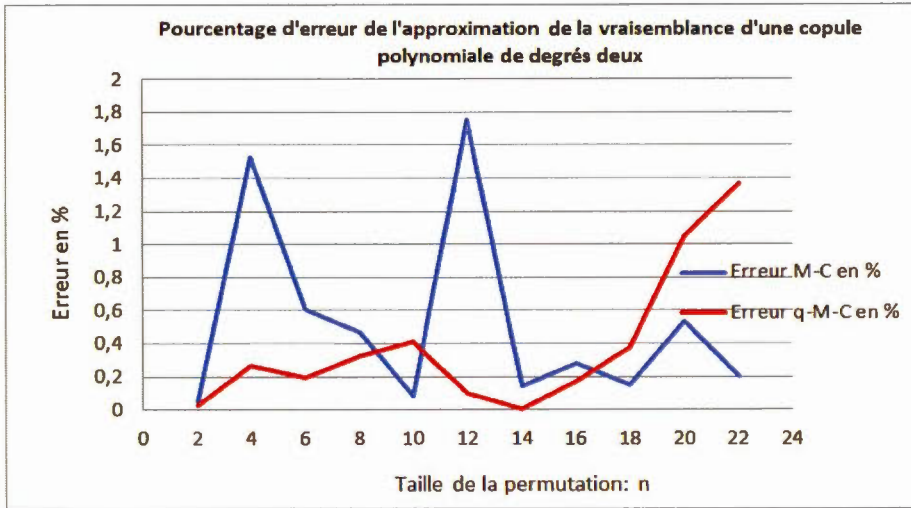


Figure 4.7 Pourcentage d'erreur de l'approximation de la vraisemblance $n!P(S(X,Y) = s)$ selon n , pour les deux méthodes, où $C_\theta(u,v) = uv(1 + \theta(1 - u)(1 - v))$, $\theta = 0,8$ et $N_{MC} = N_{qMC} = 3000$.

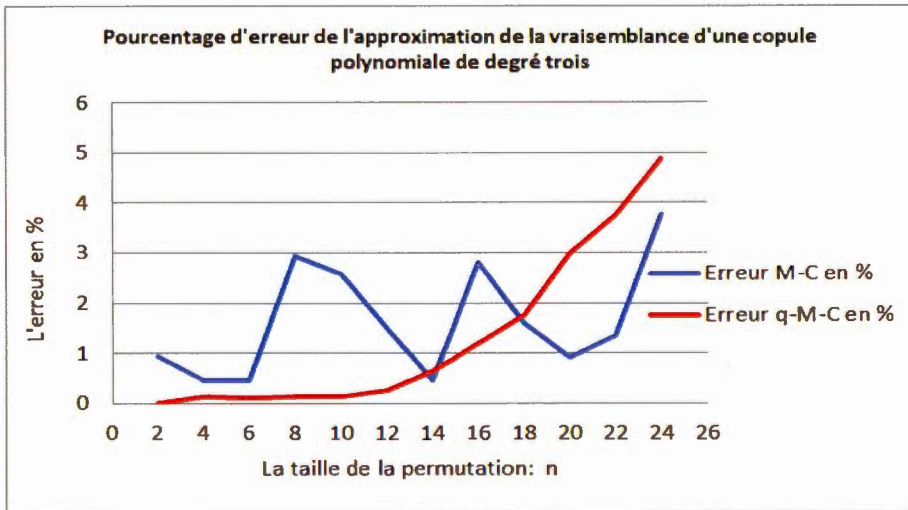


Figure 4.8 Pourcentage d'erreur de l'approximation $n!P(S(X,Y) = s)$ pour les deux méthodes, Monte-Carlo et quasi-Monte-Carlo selon n , où $C_{a,b}(u,v) = uv + uv(1-u)(1-v)[(a-b)v(1-u)+b]$, $a = -1.2$, $b = 0.6$ et $N_{MC} = N_{QMC} = 3000$.

D'après les figures 4.7 et 4.8 ci-dessus, nous remarquons que :

- Par la méthode de quasi-Monte-Carlo, les erreurs d'approximations de la vraisemblance pour des dimensions $n \leq 16$, sont relativement petites, elles ne dépassent pas 0.4% et varient très peu. En revanche, avec la méthode de Monte-Carlo, les erreurs de l'approximation varient beaucoup par rapport à n et elles ne s'approchent pas de 0% quand la taille de la permutation est relativement petite.
- Pour des dimensions, i.e. plus de 16, l'erreur de l'approximation obtenue par la méthode de quasi-Monte-Carlo augmente avec la dimension n . Cette erreur est plus significative que celle obtenue par la méthode de Monte-Carlo.
- L'approximation de la vraisemblance par la méthode de quasi-Monte-Carlo «*vr.QMC*» pour $n \leq 16$, est souvent plus proche de la valeur exacte : $n! P(S(X, Y) = s) = vr.GM(n, 0.8, 4)$. Alors qu'avec l'autre méthode, nous ne pouvons pas garantir une aussi bonne approximation que celle obtenue avec la méthode de quasi-Monte-Carlo «*vr.MC*».
- Le calcul de l'approximation de la vraisemblance pour des permutations de taille importante a été tenté, mais l'ordinateur n'a pas pu s'exécuter pour des dimensions supérieures à 24. Par contre, en augmentant significativement la taille de l'échantillon ci-haut, l'ordinateur arrive à calculer les approximations, mais nécessite beaucoup de temps.

Dans le calcul de la vraisemblance d'une copule basée sur les rangs, quand la taille de la permutation est inférieure à 16, la méthode de quasi-Monte-Carlo réussit beaucoup mieux que la méthode de Monte-Carlo, et donne une approximation assurément proche à la valeur exacte. Donc, quand la taille de la permutation est assez grande, la méthode de Monte-Carlo est relativement meilleure.

Lorsque la copule est polynomiale, la méthode de Grundmann-Möller reste évidemment la meilleure parmi les trois méthodes pour le calcul de cette vraisemblance, car elle donne sa valeur exacte.

Dans la section 4.3.2, nous avons trouvé que l'approximation de la vraisemblance par la méthode de Monte-Carlo fluctue moins et s'approche plus de la valeur exacte à mesure que la taille de l'échantillon augmente.

CONCLUSION

Le point de départ de ce mémoire était le calcul d'une approximation de la vraisemblance d'une copule basée sur les rangs, puisque sa valeur est difficile à obtenir de façon théorique. Il était donc nécessaire de calculer cette vraisemblance au moyen d'algorithmes. Nous avons exprimé cette vraisemblance à l'aide d'une espérance et d'une intégrale sur simplexe (Guillotte *et al.*, 2017). Ces deux représentations de la vraisemblance nous ont permis d'obtenir sa valeur, de façon numérique, à l'aide de trois algorithmes : Grundmann-Möller, Monte-Carlo et quasi-Monte-Carlo.

L'algorithme de Grundmann-Möller est une méthode de quadrature déterministe lourde au niveau des calculs. En général, cet algorithme n'est pas utile pour calculer la vraisemblance d'une copule. Dans ce mémoire, son importance reste pourtant significative puisqu'il calcule la valeur exacte de la vraisemblance basée sur les rangs des copules polynomiales. Il nous a donc permis de valider une comparaison numérique de deux méthodes, Monte-Carlo et quasi-Monte-Carlo, dans le calcul de la vraisemblance.

Dans la section 4.3.2, nous avons vu que l'approximation de la vraisemblance par la méthode de Monte-Carlo fluctue moins et approche davantage la valeur exacte à mesure que la taille de l'échantillon augmente.

Dans la section 4.3.3, nous avons comparé les deux méthodes en fonction de la taille de l'échantillon choisie. Avec la méthode de quasi-Monte-Carlo, nous avons vu qu'une grande taille d'échantillon n'est pas nécessairement requise (séquence de Halton) pour obtenir une meilleure approximation de la vraisemblance basée sur les rangs. La méthode de Monte-Carlo, quant à elle, requiert une très grande taille

d'échantillon pour tenter d'obtenir une bonne approximation. Par ailleurs, cette approximation fluctue beaucoup plus qu'avec la méthode de quasi-Monte-Carlo. En considérant seulement la taille d'échantillon, tout en gardant la taille du rang (permutation) relativement petite (inférieure à 12, environ), nous pourrions donc affirmer que la méthode de quasi-Monte-Carlo calcule mieux la vraisemblance.

Dans la section 4.3.4, nous avons comparé les deux méthodes en fonction de la taille du rang (permutation), tout en conservant la même taille d'échantillon. Nous avons remarqué que, pour toute taille de permutation inférieure à 16, l'erreur de l'approximation de quasi-Monte-Carlo est minime comparée à celle de Monte-Carlo. Toutefois, au-delà de 16, l'approximation de quasi-Monte-Carlo s'éloigne de plus en plus de la valeur exacte de la vraisemblance. Nous pourrions donc affirmer que la méthode de quasi-Monte-Carlo est beaucoup plus efficace quand la taille de la permutation est inférieure à 16. La méthode de Monte-Carlo, quant à elle, est probablement meilleure lorsque la taille de la permutation est supérieure à 16.

Le calcul de l'approximation de la vraisemblance, pour des permutations de taille importante, a été tenté, mais l'ordinateur n'arrive plus à calculer lorsque la taille de permutation dépasse 24. En revanche, si la taille de l'échantillon est suffisamment grande, l'ordinateur réussit à calculer l'approximation de la vraisemblance, mais l'opération est très longue.

APPENDICE A

PREUVE DE LA PROPOSITION 2.5.1

Soient (X_i, Y_i) , $i = 1, \dots, n$, des bivariées indépendantes et identiquement distribuées ayant une fonction de répartition $H \in \mathcal{H}$, avec la copule C , et leurs fonctions de répartition marginales (continues) respectivement F et G .

Soit $X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_n)$ deux vecteurs des variables aléatoires, et soit $F(X) = (F(X_1), \dots, F(X_n))$ et $G(Y) = (G(Y_1), \dots, G(Y_n))$, alors $(F(X_i), G(Y_i))$, $i = 1, \dots, n$, sont des bivariées indépendantes et identiquement distribuées dont la fonction de répartition est la copule C , et
$$P\left(R(X, Y) = R(F(X), G(Y))\right) = 1 = P\left(S(X, Y) = S(F(X), G(Y))\right).$$

Nous prenons $s \in S_n$, $E_s = \{(u, v) \in (0, 1)^{2n} : R(u, v) = (e, s)\}$,

$X = (X_1, X_2, \dots, X_n)^\top$ un vecteur aléatoire de \mathbb{R}^n

et $D_n = \{X \in \mathbb{R}^n : X_{(1)} < \dots < X_{(n)}\}$.

La distribution de $(F(X), G(Y))$, les statistiques des rangs vues au section 2.3 de chapitre II et le fait que $\left((0, 1)^2\right)^n \setminus D_n^2$ ait une mesure Lebesgue zéro, font en

sorte que :

$$\begin{aligned}
P\left(S(X, Y) = s\right) &= P\left(S(F(X), G(Y)) = s\right) \\
&= n! P\left(R(F(X), G(Y)) = (e, s)\right), \\
&= n! \int_{E_s} \prod_{i=1}^n c(u_i, v_i) du_i dv_i, \\
&= \frac{1}{n!} \int_{(0,1)^{2n}} \prod_{i=1}^n c(u_{(i)}, v_{(s(i))}) du_i dv_i, \\
&= \frac{1}{n!} E\left(\prod_{i=1}^n c\left(U_{(i)}, V_{(s(i))}\right)\right),
\end{aligned}$$

où U_1, \dots, U_n et V_1, \dots, V_n sont des variables aléatoire indépendantes et identiquement distribuées de loi uniforme sur $(0, 1)$, et $U_{(1)}, \dots, U_{(n)}$ et $V_{(1)}, \dots, V_{(n)}$ sont respectivement des statistiques d'ordre de U_1, \dots, U_n et V_1, \dots, V_n . (Guillotte *et al.*, 2017).

APPENDICE B

PREUVE DE LA PROPOSITION (2.5.2)

Pour faciliter la notation, nous notons $c(u, v) := \prod_{i=1}^n c(u_i, v_i)$, et, de la même façon, nous notons, $dudv := \prod_{i=1}^n du_i dv_i$, nous avons :

$$\begin{aligned}
 P(S(X, Y) = s) &= P(S(F(X), G(Y)) = s), \\
 &= n! P(R(F(X), G(Y)) = (e, s)), \\
 &= n! \int_{E_s} \prod_{i=1}^n c(u_i, v_i) du_i dv_i, \\
 &= n! \int_{E_s} c(u, v) dudv, \\
 &= n! \int_{\varphi_s(\Delta_n^2)} c(u, v) dudv, \\
 &= n! \int_{\Delta_n^2} c(\varphi_s(x, y)) dxdy.
 \end{aligned}$$

Cette dernière est une intégrale itérée sur $\Delta_n \times \Delta_n$. (Guillotte *et al.*, 2017).

RÉFÉRENCES

- Grundmann, A. et Möller, H.-M. (1978). Invariant integration formulas for the n -simplex by combinatorial methods. *SIAM Journal on Numerical Analysis*, 15(2), 282–290.
- Guillotte, S., Perron, F. et Segers, J. (2017). Bayesian inference based on ranks (en cours de publication).
- Halton, J. H. (1960). *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*,. Oxford University Press.
- Kuipers, L. et Niederreiter, H. (1974). *Uniform Distribution of Sequences*. John Wiley, New York.
- Nelsen, R. B. (2007). *An introduction to copulas*. Springer Science & Business Media, Portland, OR 97219-7899.
- Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*, *SIAM-CBMS Lecture Notes No.*
- Pillards, T. (2006). Quasi-monte carlo integration over a simplex and the entire space.
- Schmidt, W. M. (1972). *Irregularities of distribution VII*, *Acta Arith.* 21.
- Sklar, A. (1959). *Fonctions de répartition à n dimensions et leurs marges*,. Inst Statist Univ Paris.

Stroud, A. H. (1971). *Approximate Calculation of Multiple Integrals*. (Thèse de doctorat).

Thiémard, E. (2000). *Sur le calcul et la majoration de la discrédance à l'origine*. (Thèse de doctorat). École polytechnique fédérale de Lausanne.

van der Corput, J. G. (1935). *Verteilungsfunktionen I and II*, *Proc. Kon. Ned. Akad. v. Wetensch.* 38.